

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 14:21:26 ; Search time 181 Seconds
(without alignments)
1378.235 Million cell updates/sec

Title: US-10-082-747A-93

Perfect score: 3347

Sequence: 1 MSERGEKRGKGGKKGKGS.....QEEIQARLSSVIANODPIAV 645

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3347	100.0	645	2	Aay06635 Herugelin
2	3347	100.0	645	3	Aab36793 Human her
3	3347	100.0	645	7	Adb67619 Human her
4	3347	100.0	645	8	Adn48871 Human her
5	3347	100.0	675	2	Aar29571 Human her
6	3347	100.0	675	3	Aay71172 Human her
7	3347	100.0	675	4	Aau09887 Human her
8	3344	99.9	675	3	Aay71198 Human her
9	3344	99.9	675	3	Aay71188 Human her
10	3344	99.9	675	3	Aay71189 Human her
11	3344	99.9	675	3	Aay71191 Human her
12	3343	99.9	675	3	Aay71182 Human her
13	3343	99.9	675	3	Aay71203 Human her
14	3343	99.9	675	3	Aay71194 Human her
15	3342	99.9	675	3	Aay71180 Human her
16	3342	99.9	675	3	Aay71185 Human her
17	3342	99.9	675	3	Aay71197 Human her
18	3341	99.8	675	3	Aay71186 Human her
19	3341	99.8	675	3	Aay71183 Human her
20	3341	99.8	675	3	Aay71190 Human her
21	3341	99.8	675	3	Aay71181 Human her
22	3341	99.8	675	3	Aay71187 Human her
23	3341	99.8	675	3	Aay71201 Human her
24	3341	99.8	675	3	Aay71184 Human her
25	3340	99.8	675	3	Aay71192 Human her

26	3340	99.8	675	3	AAY71202	Aay71202 Human Her
27	3340	99.8	675	3	AAY71200	Aay71200 Human Her
28	3340	99.8	675	3	AAY71195	Aay71195 Human Her
29	3340	99.8	675	3	AAY71178	Aay71178 Human Her
30	3339	99.8	675	3	AAY71196	Aay71196 Human Her
31	3338	99.7	675	2	AAW74494	Aaw74494 Amino aci
32	3338	99.7	675	3	AAY71193	Aay71193 Human Her
33	3338	99.7	675	3	AAY71179	Aay71179 Human Her
34	3338	99.7	675	3	AAY71199	Aay71199 Human Her
35	3337	99.7	645	4	AAW67911	Aaw67911 Human NRG
36	3337	99.7	645	4	AAW67949	Aaw67949 Human NRG
37	3335	99.6	675	2	AAW74493	Aaw74493 Amino aci
38	3332	99.6	675	2	AAW74509	Aaw74509 Amino aci
39	3327	99.4	675	2	AAW74503	Aaw74503 Amino aci
40	3322	99.3	675	2	AAW74488	Aaw74488 Amino aci
41	3319	99.2	675	2	AAW74500	Aaw74500 Amino aci
42	3316	99.1	675	2	AAW74489	Aaw74489 Amino aci
43	3316	99.1	675	2	AAW74491	Aaw74491 Amino aci
44	3312	99.0	675	2	AAW74487	Aaw74487 Amino aci
45	3312	99.0	675	2	AAW74496	Aaw74496 Amino aci

ALIGNMENTS

RESULT 1
AAY06635
ID AAY06635 standard; protein; 645 AA.
XX AC AAY06635;
XX DT 26-OCT-1999 (first entry)
XX DE Herugelin-beta-1.
XX KW Hergulin-beta-1; HRG-alpha; human; ligand; HER2; HER3; HER4; receptor;
XX KW lung surfactant; respiratory distress syndrome; emphysema;
XX KW epithelial growth factor; therapy.
XX OS Homo sapiens.
XX PN WO9939729-A2.
XX PD 12-AUG-1999.
XX PF 03-FEB-1999; 99WO-US002390.
XX PR 04-FEB-1998; 98US-00020598.
XX (GETH) GENENTECH INC.
XX (IOWA) UNIV IOWA RES FOUND.
XX PI Sliwkowski M, Kern JA;
XX DR WPI; 1999-494213/41.
XX N-PSDB; AAX87701.
XX Herugelin ligands can be used to induce epithelial cell growth, and to promote repair and healing of tissue damage or injury.
XX Disclosure; Page 87-90; 120pp; English.
XX This sequence represents heregulin-beta-1 (HRG-beta-1) deduced from HRG-beta-1 cDNA (see AAX87701). The invention provides HRG ligands, including HRG-beta-1, that have affinity for and stimulate HER2, HER3 and/or HER4 receptors in autophosphorylation. A new method of treating respiratory distress syndrome in humans uses HER2, HER3 and/or HER4 receptor ligands as epithelial growth factors. A novel method of inducing epithelial cell growth and/or proliferation comprises contacting a normal epithelial cell which expresses HER2, HER3 and/or HER4 receptors with an isolated ligand which activates HER2, HER3, HER4 receptors or their combination. Also claimed are methods of increasing lung surfactant protein A, or of treating chronic obstructive pulmonary disease, respiratory distress or

CC emphysema, by administering an effective amount of an isolated HER ligand
CC to a patient
XX
SQ Sequence 645 AA;

Query Match 100.0%; Score 3347; DB 2; Length 645;
Best Local Similarity 100.0%; Pred. No. 2.2e-204;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSERKEGRGKGKKGKVERGSGKPPESAAGSQSPALPPOLKEMKQSAAGSKLVLRCTS 60
DB 1 MSERKEGRGKGKKGKVERGSGKPPESAAGSQSPALPPOLKEMKQSAAGSKLVLRCTS 60
QY 61 SEYSSLRFRKFWKNGNELNRKNPQNIKIOKKPKSELINKASLADSGEYMKVSKLGN 120
DB 61 SEYSSLRFRKFWKNGNELNRKNPQNIKIOKKPKSELINKASLADSGEYMKVSKLGN 120
QY 121 DSANANITIVESNEIITGMPASTEGAYVSESPIRISVSTEGANTSSSTSTSTTGTSHLV 180
DB 121 DSANANITIVESNEIITGMPASTEGAYVSESPIRISVSTEGANTSSSTSTSTTGTSHLV 180
QY 181 KCAEKEKTCVNGGECFVMDLSNPORYLCKPNEFTGDRCONVYMASFYKHLGIEFMEA 240
DB 181 KCAEKEKTCVNGGECFVMDLSNPORYLCKPNEFTGDRCONVYMASFYKHLGIEFMEA 240
QY 241 BELYQKRVLTITGICALLVVGIMCVVAYCKTKQKQKLDRLRQSLRSERNMMNIANG 300
DB 241 BELYQKRVLTITGICALLVVGIMCVVAYCKTKQKQKLDRLRQSLRSERNMMNIANG 300
QY 301 PHPNPPPPENVQVNVYVSKNVISSSEHIVERAEFTSFSTSHYTSTAHHSTTTTQTPSHW 360
DB 301 PHPNPPPPENVQVNVYVSKNVISSSEHIVERAEFTSFSTSHYTSTAHHSTTTTQTPSHW 360
QY 361 SNGHTESILSESHSVIWMSSVENSRHSPTGPRGLNGTGGPRECNFLRHARETPOSY 420
DB 361 SNGHTESILSESHSVIWMSSVENSRHSPTGPRGLNGTGGPRECNFLRHARETPOSY 420
QY 421 RDSPHSERVVSAMTTPARMSVDFHTPSPKSPSEMSPPVSSMTVSMPSMAVSPFMBEE 480
DB 421 RDSPHSERVVSAMTTPARMSVDFHTPSPKSPSEMSPPVSSMTVSMPSMAVSPFMBEE 480
QY 481 RPLLLVTPRLREKFDHHPQFSSFHHPNPAHDSNLSPLRIVEDEEYETTQYEPQAQ 540
DB 481 RPLLLVTPRLREKFDHHPQFSSFHHPNPAHDSNLSPLRIVEDEEYETTQYEPQAQ 540
QY 541 EPVKKLANSRRAKTKPNGHIANRLEVDNSTSSQSNSESETEDETVGDDTFFLGIQNPL 600
DB 541 EPVKKLANSRRAKTKPNGHIANRLEVDNSTSSQSNSESETEDETVGDDTFFLGIQNPL 600
QY 601 AASLEATPAFLADSRTPNAGRFSTQBEIQARLSSVIANQDPIAV 645
DB 601 AASLEATPAFLADSRTPNAGRFSTQBEIQARLSSVIANQDPIAV 645

RESULT 2

AAB36793

ID AAB36793 standard; protein; 645 AA.

XX AAB36793;

AC AAB36793;

XX 16-FEB-2001 (first entry)

DT Human heregulin-beta1 residues 175-230.

DE Hergeulin; ErbB receptor; transplantation; cancer;
XX nervous system disease; musculature; epithelium.

KW Homo sapiens.

XX US6136558-A.

XX 24-OCT-2000.

PD

XX

PP 09-FEB-1998; 98US-00020880.
XX 10-FEB-1997; 97US-0037581P.
XX (GETH) GENENTECH INC.
XX Jones JT, Fairbrother WJ, Ballinger MD, Wells JA, Sliwkowski MX;
PI WPI; 2000-678767/66.
XX New variants of heregulin, useful e.g. for treating cancer, comprises
PT specific amino acid alterations that increase affinity for ErbB
PT receptors.
XX Claim 9; Col 81-86; 58pp; English.
XX The present invention relates to variants of heregulin that can bind to
CC an ErbB receptor and include a portion of the 175-230 region of native
CC human heregulin-beta1. The variants may be used to promote ex vivo
CC survival, proliferation and differentiation of cells, particularly when
CC intended for transplantation. They may also be used to treat a wide range
CC of cancers and diseases of the nervous system, musculature and epithelium
XX
SQ Sequence 645 AA;

Query Match 100.0%; Score 3347; DB 3; Length 645;
Best Local Similarity 100.0%; Pred. No. 2.2e-204;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSERKEGRGKGKKGKVERGSGKPPESAAGSQSPALPPOLKEMKQSAAGSKLVLRCTS 60
DB 1 MSERKEGRGKGKKGKVERGSGKPPESAAGSQSPALPPOLKEMKQSAAGSKLVLRCTS 60
QY 61 SEYSSLRFRKFWKNGNELNRKNPQNIKIOKKPKSELINKASLADSGEYMKVSKLGN 120
DB 61 SEYSSLRFRKFWKNGNELNRKNPQNIKIOKKPKSELINKASLADSGEYMKVSKLGN 120
QY 121 DSANANITIVESNEIITGMPASTEGAYVSESPIRISVSTEGANTSSSTSTTGTSHLV 180
DB 121 DSANANITIVESNEIITGMPASTEGAYVSESPIRISVSTEGANTSSSTSTTGTSHLV 180
QY 181 KCAEKEKTCVNGGECFVMDLSNPORYLCKPNEFTGDRCONVYMASFYKHLGIEFMEA 240
DB 181 KCAEKEKTCVNGGECFVMDLSNPORYLCKPNEFTGDRCONVYMASFYKHLGIEFMEA 240
QY 241 BELYQKRVLTITGICALLVVGIMCVVAYCKTKQKQKLDRLRQSLRSERNMMNIANG 300
DB 241 BELYQKRVLTITGICALLVVGIMCVVAYCKTKQKQKLDRLRQSLRSERNMMNIANG 300
QY 301 PHPNPPPPENVQVNVYVSKNVISSSEHIVERAEFTSFSTSHYTSTAHHSTTTTQTPSHW 360
DB 301 PHPNPPPPENVQVNVYVSKNVISSSEHIVERAEFTSFSTSHYTSTAHHSTTTTQTPSHW 360
QY 361 SNGHTESILSESHSVIWMSSVENSRHSPTGPRGLNGTGGPRECNFLRHARETPOSY 420
DB 361 SNGHTESILSESHSVIWMSSVENSRHSPTGPRGLNGTGGPRECNFLRHARETPOSY 420
QY 421 RDSPHSERVVSAMTTPARMSVDFHTPSPKSPSEMSPPVSSMTVSMPSMAVSPFMBEE 480
DB 421 RDSPHSERVVSAMTTPARMSVDFHTPSPKSPSEMSPPVSSMTVSMPSMAVSPFMBEE 480
QY 481 RPLLLVTPRLREKFDHHPQFSSFHHPNPAHDSNLSPLRIVEDEEYETTQYEPQAQ 540
DB 481 RPLLLVTPRLREKFDHHPQFSSFHHPNPAHDSNLSPLRIVEDEEYETTQYEPQAQ 540
QY 541 EPVKKLANSRRAKTKPNGHIANRLEVDNSTSSQSNSESETEDETVGDDTFFLGIQNPL 600
DB 541 EPVKKLANSRRAKTKPNGHIANRLEVDNSTSSQSNSESETEDETVGDDTFFLGIQNPL 600
QY 601 AASLEATPAFLADSRTPNAGRFSTQBEIQARLSSVIANQDPIAV 645
DB 601 AASLEATPAFLADSRTPNAGRFSTQBEIQARLSSVIANQDPIAV 645

RESULT 3	
ADB67619	
ID	ADB67619 standard; protein; 645 AA.
XX	
AC	ADB67619;
XX	
DT	04-DEC-2003 (first entry)
XX	
DE	Human heregulin protein.
XX	
KW	cytostatic; human epidermal growth factor receptor-3; HER-3; heregulin;
KW	HER2; tyrosine kinase activity; cancer.
XX	
OS	Homo sapiens.
XX	
PN	WO2003011897-A1.
XX	
PD	13-FEB-2003.
XX	
PF	29-JUL-2002; 2002WO-US023963.
XX	
PR	27-JUL-2001; 2001US-0308341P.
XX	
PA	(REGC) UNIV CALIFORNIA.
XX	
PI	Singer E, Landgraf R, Slamon DJ, Eisenberg D;
XX	
DR	WPI; 2003-300482/29.
DR	N-PSDB; ADB67618.
XX	
PT	Novel human epidermal growth factor receptor 3 variant as agonist or
PT	antagonist of HER3 receptor, for diagnosis/treatment of cells or
PT	pathological conditions associated with aberrant expression of heregulin
PT	or HER3.
XX	
PS	Disclosure; Page 80; 137pp; English.
XX	
CC	The invention relates to a non-naturally occurring human epidermal growth
CC	factor receptor (HER)-3 variant polypeptide comprising amino acids 19-329
CC	or 20-329 of the 1342 amino acid HER3 polypeptide (ADB67617) or a
CC	sequence which differs from native HER3 polypeptide and having amino acid
CC	substitutions at residues E43, N44, K51, E64, V66 and V110 of SI, is new.
CC	The variant HER-3 specifically binds to the heregulin polypeptide
CC	(ADB67619), exhibits an impaired ability to interact with HER2
CC	polypeptide (ADB67621), or has an ability to inhibit the interaction
CC	between wild-type HER3 and heregulin. The polypeptide is useful for
CC	identifying a compound which specifically binds to heregulin binding
CC	domain in a HER3 variant polypeptide. The method further involves
CC	determining whether the test compound inhibits or enhances the heregulin
CC	induced tyrosine kinase activity associated with a HER3 polypeptide. The
CC	polypeptide is also useful for determining whether a test compound
CC	modulates the interaction between a heregulin polypeptide, and the
CC	variant HER-3 polypeptide. The HER-3 polypeptide is also useful for
CC	inhibiting the interaction between a heregulin polypeptide and HER3
CC	polypeptide, e.g. for treating cancer. The polypeptide is also useful for
CC	stimulating or activating HER3 receptor. This sequence represents the
CC	wild type human heregulin polypeptide.
XX	
SQ	Sequence 645 AA;
Query Match	100.0%; Score 3347; DB 7; Length 645;
Best Local Similarity	100.0%; Pred. No. 2.2e-204;
Matches 645; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MSERKEGRGKGKKKVERSGKKPESAGSQSPALPPQLKEMKQESAGSKVLRCETS 60
DB	1 MSERKEGRGKGKKKVERSGKKPESAGSQSPALPPQLKEMKQESAGSKVLRCETS 60
QY	61 SEYSSLRFFKWFNGNENLNKPNQIKIQKPKSELNKLADSGYMKCVSKLGN 120
DB	61 SEYSSLRFFKWFNGNENLNKPNQIKIQKPKSELNKLADSGYMKCVSKLGN 120
QY	121 DSASANITIVSNEIITGMPASTEGAYVSSSPPIRISVSTEGANTSSSTSTTTGSHLV 180
DB	121 DSASANITIVSNEIITGMPASTEGAYVSSSPPIRISVSTEGANTSSSTSTTTGSHLV 180
QY	181 KCAEKEKTFVNGGECFVWKOLSNPSRYLCKCPNEFTGDRCONYVMAFPYKHLGIEFMEA 240
DB	181 KCAEKEKTFVNGGECFVWKOLSNPSRYLCKCPNEFTGDRCONYVMAFPYKHLGIEFMEA 240
QY	241 EELYQKRVLTITIGICIALLVGIMCVAYCKTKQKQLHDLRLQSLRSERNMMNIANG 300
DB	241 EELYQKRVLTITIGICIALLVGIMCVAYCKTKQKQLHDLRLQSLRSERNMMNIANG 300
QY	301 PHHPNPPENVOLNQYVSKNVISSEHIVERAEATSFSTSHYSTAHSHTVTTQTPSHSW 360
DB	301 PHHPNPPENVOLNQYVSKNVISSEHIVERAEATSFSTSHYSTAHSHTVTTQTPSHSW 360
QY	361 SNGHTESILSSESHSVIMSSVENSRSPTGGPRGLNGTGGPRECN3FLRHARETPDSY 420
DB	361 SNGHTESILSSESHSVIMSSVENSRSPTGGPRGLNGTGGPRECN3FLRHARETPDSY 420
QY	421 RDSPHSERYVSAMTTPARMSPVDFTHTPSPKSPSPSEMSPPVSSMTVSPMAVSPFMEEE 480
DB	421 RDSPHSERYVSAMTTPARMSPVDFTHTPSPKSPSPSEMSPPVSSMTVSPMAVSPFMEEE 480
QY	481 RPLLLVTPPRLREKKFDHHPQOFSFHNPAHDSNLSPLRIVEDEBEYETTQYVEPAQ 540
DB	481 RPLLLVTPPRLREKKFDHHPQOFSFHNPAHDSNLSPLRIVEDEBEYETTQYVEPAQ 540
QY	541 EPVKLANSRRAKRTKPNGHIANRLEVDNSNTSSQSSNSESETEDESERVEDTFLGIQNPL 600
DB	541 EPVKLANSRRAKRTKPNGHIANRLEVDNSNTSSQSSNSESETEDESERVEDTFLGIQNPL 600
QY	601 AASLEATPAFLADSRTPNAGRFSTQEIQARLSSVIANQDDPIAV 645
DB	601 AASLEATPAFLADSRTPNAGRFSTQEIQARLSSVIANQDDPIAV 645
RESULT 4	
ADN48871	
ID	ADN48871 standard; protein; 645 AA.
XX	
AC	ADN48871;
DT	15-JUL-2004 (first entry)
XX	
DE	Human heregulin protein.
XX	
KW	HLF; heregulin-like factor; diagnosis; cancer; gene therapy; human.
XX	
OS	Homo sapiens.
XX	
PN	US6727077-B1.
XX	
PD	27-APR-2004.
XX	
PF	16-JUN-1998; 98US-00097681.
XX	
PR	17-JUN-1997; 97US-0049492P.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	(GEU) UNIV GEORGETOWN MEDICAL CENT.
XX	
PI	Young PE, King CR, Hijazi M, Ruben SM;
XX	
DR	WPI; 2004-338520/31.
XX	
PT	New heregulin-like factor (HLF) nucleic acid or polypeptide, useful for
PT	preparing a composition for diagnosing or treating cancer.
XX	
PS	Disclosure; SEQ ID NO 3; 48pp; English.
XX	
CC	The present invention relates to novel heregulin-like factor (HLF)
CC	polypeptides and the encoding polynucleotides. The invention is useful

XX PN WO200027426-A1.
 XX PD 18-MAY-2000.
 XX PF 28-OCT-1999; 99WO-US025744.
 XX PR 07-NOV-1998; 98US-0107522P.
 XX PA (GETH) GENENTECH INC.
 XX PI Gao W;
 XX DR WPI; 2000-376313/32.
 XX PT Method for inducing hair cell generation and inner-ear-supporting cell
 PT growth regeneration and proliferation, useful for treating hearing
 PT disorders.
 XX PS Disclosure; Page; 141pp; English.
 XX CC The patent discloses a method for inducing hair cell generation, or inner
 CC ear-supporting cell growth, regeneration, and/or proliferation, by
 CC heregulin (HRG), the ligand for HER2/HER3 dimeric receptors. Heregulin
 CC proteins function as activators of HER-2 oncogene and result from
 CC alternate splicing of a single gene mapped to chromosome 8p. The two
 CC major types, alpha and beta HRG's are based on two variant EGF-like
 CC (epidermal growth factor) domains, that differ in their C-terminal ends.
 CC This method can be used to increase the number of inner-ear-supporting
 CC cells and for treatment of hair cell related hearing disorders and
 CC disease states associated with tissue damage, e.g. ototoxic injury,
 CC acoustic assault, degenerative hearing loss, balance impairments, damage
 CC associated with surgery or physical injury and inner ear disorders
 CC related to hair cell dysfunction. The present sequence is the human
 CC heregulin ligand, HRG-beta1 protein variant-11. Note: This sequence is
 CC not found in the specification and is derived from the HRG-beta1 protein
 CC sequence found in Figure 2 (AAV71172)
 XX SQ Sequence 675 AA;
 Query Match 99.9%; Score 3344; DB 3; Length 675;
 Best Local Similarity 99.8%; Pred. No. 3.6e-204;
 Matches 644; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSERKEGRGKGGKKGKSGKPPSAAGSQSPALPPQIKEMKSOESAAGSKVLVRCETS 60
 DB 31 MSERKEGRGKGGKKGKSGKPPSAAGSQSPALPPQIKEMKSOESAAGSKVLVRCETS 90
 QY 61 SEYSLRFKFWKNGNELNKNKPNQIKIOKPKGSELINKASLADSGEYMKVSKLGN 120
 DB 91 SEYSLRFKFWKNGNELNKNKPNQIKIOKPKGSELINKASLADSGEYMKVSKLGN 150
 QY 121 DSANANITIVESNEIITGMPASTEGAYVSSSPIRISVSTEGANTSSSTSTSTTGTSHLV 180
 DB 151 DSANANITIVESNEIITGMPASTEGAYVSSSPIRISVSTEGANTSSSTSTSTTGTSHLV 210
 QY 181 KCAEKEKTCVNGGCFMVKDLSNPSRYLCKPNEFTGDRCONVWAFYKHLGIEFMEA 240
 DB 211 KCAEKEKTCVNGGCFMVKDLSNPSRYLCKPNEFTGDRCONVWAFYKHLGIEFMEA 270
 QY 241 EELYQKRVLTITGICIALLVGMVCMVAYCKTKQKQKLDRLQSLRSERNMNIANG 300
 DB 271 EELYQKRVLTITGICIALLVGMVCMVAYCKTKQKQKLDRLQSLRSERNMNIANG 330
 QY 301 PHHPNPPPNVOLNQYVSKNVISSEHIVERAEATSFSTSHVTSTAHHSTTVTQPSHSW 360
 DB 331 PHHPNPPPNVOLNQYVSKNVISSEHIVERAEATSFSTSHVTSTAHHSTTVTQPSHSW 390
 QY 361 SNGHTESILSESHVTVMSVSNSESHSPTGGPRGLNGTGGPRCNSFLRHARETPDSY 420
 DB 391 SNGHTESILSESHVTVMSVSNSESHSPTGGPRGLNGTGGPRCNSFLRHARETPDSY 450
 QY 421 RDSPHSERYVSAMTTTPARMSPVDFHTPSSPKSPPEMSPVSSMTVSPMAVSPFMEEE 480

DB 451 RDSPHSERYVSAMTTTPARMSPVDFHTPSSPKSPPEMSPVSSMTVSPMAVSPFMEEE 510
 QY 481 RPLLVTPPRLREKKFDHHPQOFSFHNPAHDSNLSPLRIVEDEEYETQSYEPAQ 540
 DB 511 RPLLVTPPRLREKKFDHHPQOFSFHNPAHDSNLSPLRIVEDEEYETQSYEPAQ 570
 QY 541 EPVKLANSRRAKRTKPNGHIANRLEVDNSTSSQSSNSESETEDERVGEDTPFLGIQNPL 600
 DB 571 EPVKLANSRRAKRTKPNGHIANRLEVDNSTSSQSSNSESETEDERVGEDTPFLGIQNPL 630
 QY 601 AASLEATPAFLADSRTPNAGRFSTQEBEQARLSSVIANQDPIAV 645
 DB 631 AASLEATPAFLADSRTPNAGRFSTQEBEQARLSSVIANQDPIAV 675
 RESULT 10
 ID AAV71189
 XX AAY71189 standard; protein; 675 AA.
 XX AC AAY71189;
 XX DT 21-SEP-2000 (first entry)
 XX DE Human Heregulin HRG-beta1 protein variant-12.
 XX KW Heregulin; HRG beta1; human; mutant; mutein; inner-ear-supporting cell;
 KW activator; HER2 receptor; HER3 receptor; HER-2 oncogene; chromosome 8p;
 KW hair cell related hearing disorder; ototoxic injury; tissue damage;
 KW acoustic assault; degenerative hearing loss; balance impairment;
 KW treatment; surgical injury; physical injury; inner ear disorder.
 XX OS Homo sapiens.
 XX OS Synthetic.
 XX FH Key
 XX FT Misc-difference 31 Location/Qualifiers
 XX FT Misc-difference 227 /label= Initiator_methionine
 FT /note= "Wild type Phe replaced with Tyr; This variable
 FT site corresponds to position 197, of the HRG-beta1 mature
 FT protein, after initiator Met (M31)"
 XX WO200027426-A1.
 XX 18-MAY-2000.
 XX 28-OCT-1999; 99WO-US025744.
 XX 07-NOV-1998; 98US-0107522P.
 XX (GETH) GENENTECH INC.
 XX Gao W;
 XX WPI; 2000-376313/32.
 XX PT Method for inducing hair cell generation and inner-ear-supporting cell
 PT growth regeneration and proliferation, useful for treating hearing
 PT disorders.
 XX Disclosure; Page; 141pp; English.
 XX CC The patent discloses a method for inducing hair cell generation, or inner
 CC ear-supporting cell growth, regeneration, and/or proliferation, by
 CC heregulin (HRG), the ligand for HER2/HER3 dimeric receptors. Heregulin
 CC proteins function as activators of HER-2 oncogene and result from
 CC alternate splicing of a single gene mapped to chromosome 8p. The two
 CC major types, alpha and beta HRG's are based on two variant EGF-like
 CC (epidermal growth factor) domains, that differ in their C-terminal ends.
 CC This method can be used to increase the number of inner-ear-supporting
 CC cells and for treatment of hair cell related hearing disorders and
 CC disease states associated with tissue damage, e.g. ototoxic injury,
 CC associated with surgery or physical injury and inner ear disorders
 CC related to hair cell dysfunction. The present sequence is the human
 CC heregulin ligand, HRG-beta1 protein variant-11. Note: This sequence is
 CC not found in the specification and is derived from the HRG-beta1 protein
 CC sequence found in Figure 2 (AAV71172)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 14:27:47 ; Search time 43 Seconds
(without alignments)
1119.736 Million cell updates/sec

Title: US-10-082-747A-93

Perfect score: 3347

Sequence: 1 MSERKEGRGKGKKRGS.....QEEIQARLSSVIANQPIAV 645

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCBUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3347	100.0	645	1	US-07-847-743B-27
2	3347	100.0	645	1	US-08-456-201-27
3	3347	100.0	645	1	US-08-428-926-4
4	3347	100.0	645	1	US-08-428-927-4
5	3347	100.0	645	1	US-08-428-298-4
6	3347	100.0	645	1	US-08-319-517-4
7	3347	100.0	645	2	US-08-456-241-27
8	3347	100.0	645	3	US-09-020-880-93
9	3347	100.0	645	3	US-09-101-544-93
10	3347	100.0	645	4	US-09-097-681-3
11	3347	100.0	645	5	PCT-US92-04295A-27
12	3347	100.0	732	1	US-07-847-743B-9
13	3347	100.0	732	1	US-08-456-201-9
14	3347	100.0	732	2	US-08-456-241-9
15	3347	100.0	732	5	PCT-US92-04295A-9
16	3343	99.9	645	3	US-08-753-007A-10
17	3343	99.9	645	3	US-09-398-496-10
18	3332.5	99.6	646	4	US-08-467-602-270
19	3332.5	99.6	646	4	US-08-411-295F-196
20	3311	98.9	669	4	US-08-467-602-273
21	3311	98.9	669	4	US-08-411-295F-199
22	3299	98.6	693	4	US-08-467-602-271
23	3299	98.6	693	4	US-08-411-295F-197
24	3286	98.2	637	4	US-08-467-602-258
25	3286	98.2	637	4	US-08-411-295F-184
26	3284	98.1	637	1	US-07-847-743B-28
27	3284	98.1	637	1	US-08-456-201-28

Query Match 100.0%; Score 3347; DB 1; Length 645;

28	3284	98.1	637	2	US-08-456-241-28	Sequence 28, Appl
29	3284	98.1	637	5	PCT-US92-04295A-28	Sequence 28, Appl
30	3277.5	97.9	716	4	US-08-467-602-274	Sequence 274, App
31	3277.5	97.9	716	4	US-08-411-295F-200	Sequence 200, App
32	3264.5	97.5	660	4	US-08-467-602-264	Sequence 264, App
33	3264.5	97.5	660	4	US-08-411-295F-190	Sequence 190, App
34	3252.5	97.2	684	4	US-08-467-602-259	Sequence 259, App
35	3252.5	97.2	684	4	US-08-411-295F-185	Sequence 185, App
36	3239	96.8	649	4	US-08-467-602-267	Sequence 267, App
37	3239	96.8	649	4	US-08-411-295F-193	Sequence 193, App
38	3231	96.5	707	4	US-08-467-602-262	Sequence 262, App
39	3231	96.5	707	4	US-08-411-295F-188	Sequence 188, App
40	3205.5	95.8	696	4	US-08-467-602-268	Sequence 268, App
41	3205.5	95.8	696	4	US-08-411-295F-194	Sequence 194, App
42	3198.5	95.6	640	4	US-08-467-602-256	Sequence 256, App
43	3198.5	95.6	640	4	US-08-411-295F-182	Sequence 182, App
44	3198	95.5	895	4	US-08-467-602-354	Sequence 354, App
45	3198	95.5	895	4	US-08-411-295F-280	Sequence 280, App

ALIGNMENTS

RESULT 1
US-07-847-743B-27
; Sequence 27, Application US/07847743B
; Patent No. 5367060
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Structure, Production and Use of
; TITLE OF INVENTION: Heregulin
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/847,743B
; FILING DATE: 19920306
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705256
; FILING DATE: 24-May-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/765212
; FILING DATE: 25-Sep-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/790801
; FILING DATE: 08-No. 5367060-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hensley, Max D.
; REGISTRATION NUMBER: 27,043
; REFERENCE/DOCKET NUMBER: 712P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1489
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 645 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-847-743B-27

Best Local Similarity 100.0%; Pred. No. 1.3e-244;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSERKEGRGKGKGGKGGKPPESAAQSOSPALPOLKEMKQESAAAGSKLVLRCTS 60
DB 1 MSERKEGRGKGKGGKGGKGGKPPESAAQSOSPALPOLKEMKQESAAAGSKLVLRCTS 60

QY 61 SEYSSLRFFKFWKNGNELNKNPQNIKIQQKPKSELINKASLADSGEYMKVSKLGN 120
DB 61 SEYSSLRFFKFWKNGNELNKNPQNIKIQQKPKSELINKASLADSGEYMKVSKLGN 120

QY 121 DSANITIVESNEIITGMPASTEGAYVSSSPIRISVSTEGANTSSSTSTTTGTSHLV 180
DB 121 DSANITIVESNEIITGMPASTEGAYVSSSPIRISVSTEGANTSSSTSTTTGTSHLV 180

QY 181 KCAEKEKTCVNGGECFVMDLSNPSRYLCKPNEFTGDRCONVYMASFYKHLGIEFMEA 240
DB 181 KCAEKEKTCVNGGECFVMDLSNPSRYLCKPNEFTGDRCONVYMASFYKHLGIEFMEA 240

QY 241 EELYQKRVLTITGICALLVVGIMCVAYCKTKQKSELINKASLADSGEYMKVSKLGN 300
DB 241 EELYQKRVLTITGICALLVVGIMCVAYCKTKQKSELINKASLADSGEYMKVSKLGN 300

QY 301 PHHPNPPENVOLVNOYVSKNVISSHIVVEREATSFSTSHYTSTAHSSTTTTQPSHW 360
DB 301 PHHPNPPENVOLVNOYVSKNVISSHIVVEREATSFSTSHYTSTAHSSTTTTQPSHW 360

QY 361 SNGHTESILSESHSVIWMSSVNSRHSSPTGGPRGLNGTGGPRECNFLRHARETDPDSY 420
DB 361 SNGHTESILSESHSVIWMSSVNSRHSSPTGGPRGLNGTGGPRECNFLRHARETDPDSY 420

QY 421 RDSPHSERVVSAMTTPARMSPVDFHTPSPKPPSEMSPVSSMTVMPMAVSPFMBEE 480
DB 421 RDSPHSERVVSAMTTPARMSPVDFHTPSPKPPSEMSPVSSMTVMPMAVSPFMBEE 480

QY 481 RPLLLVTPRLREKFKDHPHQFSSPHHPNPAHDSNLSLPASPLRIVEDEEYETQYEPQA 540
DB 481 RPLLLVTPRLREKFKDHPHQFSSPHHPNPAHDSNLSLPASPLRIVEDEEYETQYEPQA 540

QY 541 EPVKLANSRRAKTKPNGHIANRLEVDNSTSSQSNSESETEDERVGEDTFFLGIQNPL 600
DB 541 EPVKLANSRRAKTKPNGHIANRLEVDNSTSSQSNSESETEDERVGEDTFFLGIQNPL 600

QY 601 AASLEATPAFLADSRTPNAGRFSTQBEIQARLSSVIANQDPIAV 645
DB 601 AASLEATPAFLADSRTPNAGRFSTQBEIQARLSSVIANQDPIAV 645

RESULT 2

US-08-456-201-27
Sequence 27, Application US/08456201
Patent No. 5641869

GENERAL INFORMATION:
APPLICANT: Vandlen, Richard L.,
APPLICANT: Holmes, William E.
TITLE OF INVENTION: Structure, Production and Use of
TITLE OF INVENTION: Heregulin 2 ligands
NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,201
FILING DATE:

CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/126,145
FILING DATE:
APPLICATION NUMBER: 07/880,917
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/705256
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/765212
APPLICATION NUMBER: 07/790801
ATTORNEY/AGENT INFORMATION:
NAME: Hensley, Max D.
REGISTRATION NUMBER: 27,034
REFERENCE/DOCKET NUMBER: 712P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/266-1489
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 645 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-456-201-27

Query Match 100.0%; Score 3347; DB 1; Length 645;
Best Local Similarity 100.0%; Pred. No. 1.3e-244;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSERKEGRGKGKGGKGGKPPESAAQSOSPALPOLKEMKQESAAAGSKLVLRCTS 60
DB 1 MSERKEGRGKGKGGKGGKGGKPPESAAQSOSPALPOLKEMKQESAAAGSKLVLRCTS 60

QY 61 SEYSSLRFFKFWKNGNELNKNPQNIKIQQKPKSELINKASLADSGEYMKVSKLGN 120
DB 61 SEYSSLRFFKFWKNGNELNKNPQNIKIQQKPKSELINKASLADSGEYMKVSKLGN 120

QY 121 DSANITIVESNEIITGMPASTEGAYVSSSPIRISVSTEGANTSSSTSTTTGTSHLV 180
DB 121 DSANITIVESNEIITGMPASTEGAYVSSSPIRISVSTEGANTSSSTSTTTGTSHLV 180

QY 181 KCAEKEKTCVNGGECFVMDLSNPSRYLCKPNEFTGDRCONVYMASFYKHLGIEFMEA 240
DB 181 KCAEKEKTCVNGGECFVMDLSNPSRYLCKPNEFTGDRCONVYMASFYKHLGIEFMEA 240

QY 241 EELYQKRVLTITGICALLVVGIMCVAYCKTKQKSELINKASLADSGEYMKVSKLGN 300
DB 241 EELYQKRVLTITGICALLVVGIMCVAYCKTKQKSELINKASLADSGEYMKVSKLGN 300

QY 301 PHHPNPPENVOLVNOYVSKNVISSHIVVEREATSFSTSHYTSTAHSSTTTTQPSHW 360
DB 301 PHHPNPPENVOLVNOYVSKNVISSHIVVEREATSFSTSHYTSTAHSSTTTTQPSHW 360

QY 361 SNGHTESILSESHSVIWMSSVNSRHSSPTGGPRGLNGTGGPRECNFLRHARETDPDSY 420
DB 361 SNGHTESILSESHSVIWMSSVNSRHSSPTGGPRGLNGTGGPRECNFLRHARETDPDSY 420

QY 421 RDSPHSERVVSAMTTPARMSPVDFHTPSPKPPSEMSPVSSMTVMPMAVSPFMBEE 480
DB 421 RDSPHSERVVSAMTTPARMSPVDFHTPSPKPPSEMSPVSSMTVMPMAVSPFMBEE 480

QY 481 RPLLLVTPRLREKFKDHPHQFSSPHHPNPAHDSNLSLPASPLRIVEDEEYETQYEPQA 540
DB 481 RPLLLVTPRLREKFKDHPHQFSSPHHPNPAHDSNLSLPASPLRIVEDEEYETQYEPQA 540

QY 541 EPVKLANSRRAKTKPNGHIANRLEVDNSTSSQSNSESETEDERVGEDTFFLGIQNPL 600
DB 541 EPVKLANSRRAKTKPNGHIANRLEVDNSTSSQSNSESETEDERVGEDTFFLGIQNPL 600

QY 601 AASLEATPAFLADSRTPNAGRFSTQBEIQARLSSVIANQDPIAV 645
DB 601 AASLEATPAFLADSRTPNAGRFSTQBEIQARLSSVIANQDPIAV 645


```
Db 61 SEYSSLRFKWFGKNGNELNKNKPNQIKIQQKPKSELKINKASLADSGEYMKVLSKLN 120
Qy 121 DSANITIVESNEIITGMPASTEGAYVSESPRISVSTEGANTSSSTSTTGTSHLV 180
Db 121 DSANITIVESNEIITGMPASTEGAYVSESPRISVSTEGANTSSSTSTTGTSHLV 180
Qy 181 KCAEKEKTFVNGGECFVMDLSNPSRYLCKPNEFTGDRCONVYMASFYKHLGIEFMEA 240
Db 181 KCAEKEKTFVNGGECFVMDLSNPSRYLCKPNEFTGDRCONVYMASFYKHLGIEFMEA 240
Qy 241 EELYQKRVLTITGICIALLVGIMCVVAYCKTKQKRLHDLRQSLRSERNMMNIANG 300
Db 241 EELYQKRVLTITGICIALLVGIMCVVAYCKTKQKRLHDLRQSLRSERNMMNIANG 300
Qy 301 PHPNPPPPNVQVNVYVKNVISEHIVERAEATSFSTSHYTAHSHSTTTVQTPSHW 360
Db 301 PHPNPPPPNVQVNVYVKNVISEHIVERAEATSFSTSHYTAHSHSTTTVQTPSHW 360
Qy 361 SNGHTESILSESHSVIWMSSVNSRHSSPTGGPRGLNGTGGPRECNSFLRHARETPOSY 420
Db 361 SNGHTESILSESHSVIWMSSVNSRHSSPTGGPRGLNGTGGPRECNSFLRHARETPOSY 420
Qy 421 RDSPHSERVVSAMTTPARMSPVDFHTPPSPKPPSPSEMSPPVSSMTVSPMSMAVSPFMBEE 480
Db 421 RDSPHSERVVSAMTTPARMSPVDFHTPPSPKPPSPSEMSPPVSSMTVSPMSMAVSPFMBEE 480
Qy 481 RPLLLVTPRLREKFKDHPHQFSSPHHNPADNSLNPASPLRIVEDEYEYETQYEPQAQ 540
Db 481 RPLLLVTPRLREKFKDHPHQFSSPHHNPADNSLNPASPLRIVEDEYEYETQYEPQAQ 540
Qy 541 EPVKLANRRARAKTKPNGHIANRLEVDNSTSSQSNSESETEDERVGDTFPLGIQNPL 600
Db 541 EPVKLANRRARAKTKPNGHIANRLEVDNSTSSQSNSESETEDERVGDTFPLGIQNPL 600
Qy 601 AASLEATPAFLADSRTPNAGRFSTQEIQARLSSVIANQDPIAV 645
Db 601 AASLEATPAFLADSRTPNAGRFSTQEIQARLSSVIANQDPIAV 645
```

RESULT 5

```
US-08-428-298-4
; Sequence 4, Application US/08428298
; Patent No. 5763213
; GENERAL INFORMATION:
; APPLICANT: Ho, Wei-Hsien
; APPLICANT: Osherooff, Phyllis L.
; TITLE OF INVENTION: SENSORY AND MOTOR NEURON DERIVED FACTOR (SMDF)
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,298
; FILING DATE: 25-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/339517
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 853D2
; TELECOMMUNICATION INFORMATION:
```

RESULT 6

```
US-08-339-517-4
; Sequence 4, Application US/08339517
; Patent No. 5770567
; GENERAL INFORMATION:
; APPLICANT: Ho, Wei-Hsien
; APPLICANT: Osherooff, Phyllis L.
; TITLE OF INVENTION: SENSORY AND MOTOR NEURON DERIVED FACTOR (SMDF)
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
```

```
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 645 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-428-298-4
```

```
Query Match 100.0%; Score 3347; DB 1; Length 645;
Best Local Similarity 100.0%; Pred. No. 1.3e-244;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSERKEGKGKGGKKGSGKSPESAAGSQSPALPPOLKEMKQESAGSKLVLRCTS 60
Db 1 MSERKEGKGKGGKKGSGKSPESAAGSQSPALPPOLKEMKQESAGSKLVLRCTS 60
Qy 61 SEYSSLRFKWFGKNGNELNKNKPNQIKIQQKPKSELKINKASLADSGEYMKVLSKLN 120
Db 61 SEYSSLRFKWFGKNGNELNKNKPNQIKIQQKPKSELKINKASLADSGEYMKVLSKLN 120
Qy 121 DSANITIVESNEIITGMPASTEGAYVSESPRISVSTEGANTSSSTSTTGTSHLV 180
Db 121 DSANITIVESNEIITGMPASTEGAYVSESPRISVSTEGANTSSSTSTTGTSHLV 180
Qy 181 KCAEKEKTFVNGGECFVMDLSNPSRYLCKPNEFTGDRCONVYMASFYKHLGIEFMEA 240
Db 181 KCAEKEKTFVNGGECFVMDLSNPSRYLCKPNEFTGDRCONVYMASFYKHLGIEFMEA 240
Qy 241 EELYQKRVLTITGICIALLVGIMCVVAYCKTKQKRLHDLRQSLRSERNMMNIANG 300
Db 241 EELYQKRVLTITGICIALLVGIMCVVAYCKTKQKRLHDLRQSLRSERNMMNIANG 300
Qy 301 PHPNPPPPNVQVNVYVKNVISEHIVERAEATSFSTSHYTAHSHSTTTVQTPSHW 360
Db 301 PHPNPPPPNVQVNVYVKNVISEHIVERAEATSFSTSHYTAHSHSTTTVQTPSHW 360
Qy 361 SNGHTESILSESHSVIWMSSVNSRHSSPTGGPRGLNGTGGPRECNSFLRHARETPOSY 420
Db 361 SNGHTESILSESHSVIWMSSVNSRHSSPTGGPRGLNGTGGPRECNSFLRHARETPOSY 420
Qy 421 RDSPHSERVVSAMTTPARMSPVDFHTPPSPKPPSPSEMSPPVSSMTVSPMSMAVSPFMBEE 480
Db 421 RDSPHSERVVSAMTTPARMSPVDFHTPPSPKPPSPSEMSPPVSSMTVSPMSMAVSPFMBEE 480
Qy 481 RPLLLVTPRLREKFKDHPHQFSSPHHNPADNSLNPASPLRIVEDEYEYETQYEPQAQ 540
Db 481 RPLLLVTPRLREKFKDHPHQFSSPHHNPADNSLNPASPLRIVEDEYEYETQYEPQAQ 540
Qy 541 EPVKLANRRARAKTKPNGHIANRLEVDNSTSSQSNSESETEDERVGDTFPLGIQNPL 600
Db 541 EPVKLANRRARAKTKPNGHIANRLEVDNSTSSQSNSESETEDERVGDTFPLGIQNPL 600
Qy 601 AASLEATPAFLADSRTPNAGRFSTQEIQARLSSVIANQDPIAV 645
Db 601 AASLEATPAFLADSRTPNAGRFSTQEIQARLSSVIANQDPIAV 645
```



```
Db 61 SEYSSLRKFWKNGNELNRKNKPNQIKIQKPKGSELINKASLADSGEYMKCVSKLGN 120
QY 121 DSANANITIVESNEIITGMPASTEGAYVSESPIRISVSTEGANTSSSTSTSTTGTSHLV 180
Db 121 DSANANITIVESNEIITGMPASTEGAYVSESPIRISVSTEGANTSSSTSTSTTGTSHLV 180
QY 181 KCAEKEKTCVNGGECFVWKDLNPSRYLCKPCNEFTGDRCONYVNASFYKHLGIEFMEA 240
Db 181 KCAEKEKTCVNGGECFVWKDLNPSRYLCKPCNEFTGDRCONYVNASFYKHLGIEFMEA 240
QY 241 EELYQKRVLTITGICIAALLVVGIMCVVAYCKTKQKQKLDHRLRQSLRSENNMNIANG 300
Db 241 EELYQKRVLTITGICIAALLVVGIMCVVAYCKTKQKQKLDHRLRQSLRSENNMNIANG 300
QY 301 PHPNPPPPENQVQVYVKNVISSEHIVERAEATSFSTSHYTTSTAHHSTTTVQTPSHSW 360
Db 301 PHPNPPPPENQVQVYVKNVISSEHIVERAEATSFSTSHYTTSTAHHSTTTVQTPSHSW 360
QY 361 SNGHTESILSESHVIMSVSNRHSPTGGPRGLNGTGGPRECNSFLRHARETPDSY 420
Db 361 SNGHTESILSESHVIMSVSNRHSPTGGPRGLNGTGGPRECNSFLRHARETPDSY 420
QY 421 RDSPHSERVYSAMTTTPARMSPVDFHTPSSPKSPPEMSPPVSSMTVSMPSMAVSPFMEEE 480
Db 421 RDSPHSERVYSAMTTTPARMSPVDFHTPSSPKSPPEMSPPVSSMTVSMPSMAVSPFMEEE 480
QY 481 RPLLLVTPRLREKFDHHPQOFSFHHNPAHDSNLSPLASPLRIVEDEYEYETTOYEPAQ 540
Db 481 RPLLLVTPRLREKFDHHPQOFSFHHNPAHDSNLSPLASPLRIVEDEYEYETTOYEPAQ 540
QY 541 EPVKLANRRAKRTKPNGHIANRLEVDNSTSSQSSNSESETEDERVGDDTFFLGIQNPL 600
Db 541 EPVKLANRRAKRTKPNGHIANRLEVDNSTSSQSSNSESETEDERVGDDTFFLGIQNPL 600
QY 601 AASLEATPAFLADSRTPNAGRFSTQEEIQARLSSVIANQDPIAV 645
Db 601 AASLEATPAFLADSRTPNAGRFSTQEEIQARLSSVIANQDPIAV 645

RESULT 8
US-09-020-880-93
; Sequence 93, Application US/09020880A
; Patent No. 6136558
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ballinger, Marcus D.
; APPLICANT: Jones, Jennifer T.
; APPLICANT: Fairbrother, Wayne J.
; APPLICANT: Sliwowski, Mark X.
; APPLICANT: Wells, James A.
; TITLE OF INVENTION: HERGULIN VARIANTS
; FILE REFERENCE: 14918-720CON1
; CURRENT APPLICATION NUMBER: US/09/020,880A
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: US 60/037,581
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 93
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-020-880-93

Query Match 100.0%; Score 3347; DB 3; Length 645;
Best Local Similarity 100.0%; Pred. No. 1.3e-244;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MSERKEGRGKGKKGKRGSGKGPESAGSQSPALPQPKEMKQSAAGSKLVLCRTS 60
Db 1 MSERKEGRGKGKKGKRGSGKGPESAGSQSPALPQPKEMKQSAAGSKLVLCRTS 60
```

```
QY 61 SEYSSLRKFWKNGNELNRKNKPNQIKIQKPKGSELINKASLADSGEYMKCVSKLGN 120
Db 61 SEYSSLRKFWKNGNELNRKNKPNQIKIQKPKGSELINKASLADSGEYMKCVSKLGN 120
QY 121 DSANANITIVESNEIITGMPASTEGAYVSESPIRISVSTEGANTSSSTSTSTTGTSHLV 180
Db 121 DSANANITIVESNEIITGMPASTEGAYVSESPIRISVSTEGANTSSSTSTSTTGTSHLV 180
QY 181 KCAEKEKTCVNGGECFVWKDLNPSRYLCKPCNEFTGDRCONYVNASFYKHLGIEFMEA 240
Db 181 KCAEKEKTCVNGGECFVWKDLNPSRYLCKPCNEFTGDRCONYVNASFYKHLGIEFMEA 240
QY 241 EELYQKRVLTITGICIAALLVVGIMCVVAYCKTKQKQKLDHRLRQSLRSENNMNIANG 300
Db 241 EELYQKRVLTITGICIAALLVVGIMCVVAYCKTKQKQKLDHRLRQSLRSENNMNIANG 300
QY 301 PHPNPPPPENQVQVYVKNVISSEHIVERAEATSFSTSHYTTSTAHHSTTTVQTPSHSW 360
Db 301 PHPNPPPPENQVQVYVKNVISSEHIVERAEATSFSTSHYTTSTAHHSTTTVQTPSHSW 360
QY 361 SNGHTESILSESHVIMSVSNRHSPTGGPRGLNGTGGPRECNSFLRHARETPDSY 420
Db 361 SNGHTESILSESHVIMSVSNRHSPTGGPRGLNGTGGPRECNSFLRHARETPDSY 420
QY 421 RDSPHSERVYSAMTTTPARMSPVDFHTPSSPKSPPEMSPPVSSMTVSMPSMAVSPFMEEE 480
Db 421 RDSPHSERVYSAMTTTPARMSPVDFHTPSSPKSPPEMSPPVSSMTVSMPSMAVSPFMEEE 480
QY 481 RPLLLVTPRLREKFDHHPQOFSFHHNPAHDSNLSPLASPLRIVEDEYEYETTOYEPAQ 540
Db 481 RPLLLVTPRLREKFDHHPQOFSFHHNPAHDSNLSPLASPLRIVEDEYEYETTOYEPAQ 540
QY 541 EPVKLANRRAKRTKPNGHIANRLEVDNSTSSQSSNSESETEDERVGDDTFFLGIQNPL 600
Db 541 EPVKLANRRAKRTKPNGHIANRLEVDNSTSSQSSNSESETEDERVGDDTFFLGIQNPL 600
QY 601 AASLEATPAFLADSRTPNAGRFSTQEEIQARLSSVIANQDPIAV 645
Db 601 AASLEATPAFLADSRTPNAGRFSTQEEIQARLSSVIANQDPIAV 645
```

```
RESULT 9
US-09-101-544-93
; Sequence 93, Application US/09101544
; Patent No. 6387638
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ballinger, Marcus D.
; APPLICANT: Jones, Jennifer T.
; APPLICANT: Fairbrother, Wayne J.
; APPLICANT: Sliwowski, Mark X.
; APPLICANT: Wells, James A.
; TITLE OF INVENTION: HERGULIN VARIANTS
; FILE REFERENCE: 14918-720CON2
; CURRENT APPLICATION NUMBER: US/09/101,544
; CURRENT FILING DATE: 1998-07-17
; PRIOR APPLICATION NUMBER: US 09/020,880
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/037,581
; PRIOR FILING DATE: 1997-02-10
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-101-544-93
```

```
Query Match 100.0%; Score 3347; DB 3; Length 645;
Best Local Similarity 100.0%; Pred. No. 1.3e-244;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSERKEGRGKGKKGKRGSGKGPESAGSQSPALPQPKEMKQSAAGSKLVLCRTS 60
```


US-08-456-241-9
; Sequence 9, Application US/08456241
; Patent No. 5840525
; GENERAL INFORMATION:
; APPLICANT: Vanden, Richard L.
; APPLICANT: Holmes, William E.
; TITLE OF INVENTION: STRUCTURE, PRODUCTION AND USE OF
; TITLE OF INVENTION: HEREGULIN
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,241
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/126145
; FILING DATE: 23-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/880917
; FILING DATE: 11-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/847743
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/790801
; FILING DATE: 08-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/765212
; FILING DATE: 25-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705256
; FILING DATE: 24-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 712P4CID1
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 732 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-456-241-9

Query Match 100.0%; Score 3347; DB 2; Length 732;
Best Local Similarity 100.0%; Pred. NO. 1.6e-244;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSERKEGRGKGKKGKPSAAGSQSPALPPOLKEMKQESAAGSKLVLCRTS 60
DB 31 MSERKEGRGKGKKGKPSAAGSQSPALPPOLKEMKQESAAGSKLVLCRTS 90
QY 61 SYSSSLRFKFKNGNELNKNPQNIKIQKPGKSELINKASLADSGEYMKVTSKLG 120
DB 91 SYSSSLRFKFKNGNELNKNPQNIKIQKPGKSELINKASLADSGEYMKVTSKLG 150
QY 121 DSANANITIVESNEITGMIPASTGAYVSSPIRISVSTEGANTSSSTSTSTGTHLV 180
DB 151 DSANANITIVESNEITGMIPASTGAYVSSPIRISVSTEGANTSSSTSTSTGTHLV 210

QY 181 KCAEKEKTFVNGGECFVWKDLSNPRLCKPCNEFTGDRCONVYMASFYKHLGIEFMEA 240
DB 211 KCAEKEKTFVNGGECFVWKDLSNPRLCKPCNEFTGDRCONVYMASFYKHLGIEFMEA 270
QY 241 EELYQKRVLTITIGICIALLVGIMCVVAYCKTKQKQKLDRLQSLRSENNMMNIANG 300
DB 271 EELYQKRVLTITIGICIALLVGIMCVVAYCKTKQKQKLDRLQSLRSENNMMNIANG 330
QY 301 PHHPNPPENQVQVYKQVISEHIVERAEATSFSTSHYTTAHHSTTTVTPSHSW 360
DB 331 PHHPNPPENQVQVYKQVISEHIVERAEATSFSTSHYTTAHHSTTTVTPSHSW 390
QY 361 SNGHTESILSSESHSVIWMSSVENSRRHSPTGGPGRGLNGTGPGRECSFLRHARETPDSY 420
DB 391 SNGHTESILSSESHSVIWMSSVENSRRHSPTGGPGRGLNGTGPGRECSFLRHARETPDSY 450
QY 421 RDSPHSERVVSAMTTPARMSFVDFHTFSSPKSPFSEMSPPVSSMTVSMPSMAVSPFMEEE 480
DB 451 RDSPHSERVVSAMTTPARMSFVDFHTFSSPKSPFSEMSPPVSSMTVSMPSMAVSPFMEEE 510
QY 481 RPLLLVTPPRLREKKFDHHPQFSSFHHPNPAHDSNLSLPASPLRIVEDEEYETTQYEPAQ 540
DB 511 RPLLLVTPPRLREKKFDHHPQFSSFHHPNPAHDSNLSLPASPLRIVEDEEYETTQYEPAQ 570
QY 541 EPVKKLANSRRAKTKPNGHIANRLEVDNSNTSSOSSNSESETEDERVGEDTFFLGIONPL 600
DB 571 EPVKKLANSRRAKTKPNGHIANRLEVDNSNTSSOSSNSESETEDERVGEDTFFLGIONPL 630
QY 601 AASLEATPAFLADSRTPNAGRFSTQBEIQARLSSVIANQDPIAV 645
DB 631 AASLEATPAFLADSRTPNAGRFSTQBEIQARLSSVIANQDPIAV 675

RESULT 15
PCT-US92-04295A-9
; Sequence 9, Application PC/TUS9204295A
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Structure, Production and Use of
; TITLE OF INVENTION: Heregulin
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/04295A
; FILING DATE: 19920521
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705256
; FILING DATE: 24-MAY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/765212
; FILING DATE: 25-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/790801
; FILING DATE: 08-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/847743
; FILING DATE: 06-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hensley, Max D.
; REGISTRATION NUMBER: 27,043

```
; REFERENCE/DOCKET NUMBER: 712P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 732 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
PCT-US92-04295A-9

Query Match      100.0%; Score 3347; DB 5; Length 732;
Best Local Similarity 100.0%; Pred. No. 1.6e-244;
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSERKEGRGKGKKGKESGKPEAAGSQSPALPOLKEMKSOESAAGSKLVLRCEYS 60
Db 31 MSERKEGRGKGKKGKESGKPEAAGSQSPALPOLKEMKSOESAAGSKLVLRCEYS 90

QY 61 SEYSSLRFKWFKNGNELNRKPKQNIKIQQKPKSELIRINKASLADSGEYMKCVISKLG 120
Db 91 SEYSSLRFKWFKNGNELNRKPKQNIKIQQKPKSELIRINKASLADSGEYMKCVISKLG 150

QY 121 DSANITIVESNEIITGMPASTEGAYVSESPIRISVSTEGANTSSSTSTSTGTSHLV 180
Db 151 DSANITIVESNEIITGMPASTEGAYVSESPIRISVSTEGANTSSSTSTSTGTSHLV 210

QY 181 KCAEKEKTCVNGGECFMVKDLSNPSRYLCKNEFTGDRCONYMASFYKHLGIEFMEA 240
Db 211 KCAEKEKTCVNGGECFMVKDLSNPSRYLCKNEFTGDRCONYMASFYKHLGIEFMEA 270

QY 241 EELYQKRVLTITIGICALLVVGIMCVVAYCKTKQKQKLDRLRQSLRSENNMMNIANG 300
Db 271 EELYQKRVLTITIGICALLVVGIMCVVAYCKTKQKQKLDRLRQSLRSENNMMNIANG 330

QY 301 PHPNPPPNVQLVQVSKNVISSEHIVERAEATSFSTSHYTSTAHHSTTVTQTPSHSW 360
Db 331 PHPNPPPNVQLVQVSKNVISSEHIVERAEATSFSTSHYTSTAHHSTTVTQTPSHSW 390

QY 361 SNGHTSIISESHSVIWMSSVENSRRHSSPTGGPRGLNGTGGPRECNFLRHARETPDSY 420
Db 391 SNGHTSIISESHSVIWMSSVENSRRHSSPTGGPRGLNGTGGPRECNFLRHARETPDSY 450

QY 421 RDSPHSERVVSAMTTTPARMSPVDFHTPSGPKSPSEMSPPVSSMTVSMPSMAVSPMEEE 480
Db 451 RDSPHSERVVSAMTTTPARMSPVDFHTPSGPKSPSEMSPPVSSMTVSMPSMAVSPMEEE 510

QY 481 RPLLLVTPRLREKFDHHPQFSSFHHNPAHDSNLSPLASPLAIVEDEEYETTQYEYPAQ 540
Db 511 RPLLLVTPRLREKFDHHPQFSSFHHNPAHDSNLSPLASPLAIVEDEEYETTQYEYPAQ 570

QY 541 EPVKKLANSERAKRTYPNGHIANRLEVDNSTSSQSSNSETEDEYVGEDTFFLGIQNPL 600
Db 571 EPVKKLANSERAKRTYPNGHIANRLEVDNSTSSQSSNSETEDEYVGEDTFFLGIQNPL 630

QY 601 AASLEATPAFLADSRTPNAGRFSTQEEIQARLSSVIANQDPIAV 645
Db 631 AASLEATPAFLADSRTPNAGRFSTQEEIQARLSSVIANQDPIAV 675
```

Search completed: April 12, 2005, 14:43:38
Job time : 45 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 14:22:26 ; Search time 184 Seconds
(without alignments)
1795.060 Million cell updates/sec

Title: US-10-082-747A-93
Perfect score: 3347
Sequence: 1 MSERKEGRGKGKKKRGSS.....QBEIQARLSSVIANQDPIAV 645

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1612378 seqs, 512079187 residues
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3337	99.7	645	Q7RTW4	Q7rtw4 homo sapien
2	3280	98.0	637	Q7RTW3	Q7rtw3 homo sapien
3	3193.5	95.4	639	1 NRGI_HUMAN	Q02297 h pro-neure
4	3192.5	95.4	640	Q7RTV8	Q7rtv8 homo sapien
5	3149	94.1	645	Q6DR98	Q6dr98 mus musculu
6	3058.5	91.4	662	1 NRGI_RAT	P43322 r pro-neure
7	3002	89.7	782	Q9ESA5	Q9esa5 rattus norv
8	2526.5	75.5	602	1 NRGI_CHICK	Q05199 gallus gall
9	2395	71.6	700	Q6DR99	Q6dr99 mus musculu
10	2354.5	70.3	700	Q9ESB1	Q9esb1 rattus norv
11	2211	66.1	695	Q9ESB0	Q9esb0 rattus norv
12	2102	62.8	677	1 NRGI_XENLA	Q93383 xenopus lae
13	2080.5	62.2	462	Q7RTW1	Q7rtw1 homo sapien
14	1916	57.2	461	Q35947	Q35947 mesocricetu
15	1875.5	56.0	394	Q6TGK9	Q6tgk9 oryctolagus
16	1187	35.5	241	Q6PK61	Q6pk61 homo sapien
17	1193	35.3	241	Q7RTW0	Q7rtw0 homo sapien
18	1176	35.1	298	Q9ESA9	Q9esa9 rattus norv
19	1098	32.8	241	Q07112	Q07112 bos taurus
20	1071	32.0	211	Q7RTW5	Q7rtw5 homo sapien
21	1058	31.6	756	1 NRG2_MOUSE	P56974 mus musculu
22	1028	30.7	868	1 NRG2_RAT	Q35569 rattus norv
23	1003.5	30.0	317	Q9ESA3	Q9esa3 rattus norv
24	996.5	29.8	342	Q9ESAL	Q9esal rattus norv
25	993	29.7	850	1 NRG2_HUMAN	O14511 mus sapien
26	991	29.6	211	Q8BK18	Q8bk18 mus musculu
27	987.5	29.5	323	Q9ESA2	Q9esa2 rattus norv
28	833	24.9	167	Q8NFN2	Q8nfn2 homo sapien
29	823	24.6	422	Q7RTV9	Q7rtv9 homo sapien
30	817.5	24.4	167	Q862K0	Q862k0 bos taurus
31	503	15.0	348	Q8NFN3	Q8nfn3 homo sapien

RESULT 1

ID	Q7RTW4	PRELIMINARY;	PRT;	645 AA.
AC	Q7RTW4;			
DT	01-MAR-2004 (TrEMBLrel. 26, Created)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Neuregulin 1 isoform HRG-beta1.			
GN	Name=NRG1;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RX	PubMed=12145742;			
RA	Steinsson H., Sigurdsson E., Steinthorsdottir V., Bjornsdottir S.,			
RA	Sigmundsson T., Ghosh S., Brynjolfsson J., Gunnarsdottir S.,			
RA	Ivarsson O., Chou T.T., Hjaltason O., Birgisdottir B., Jonsson H.,			
RA	Gudnadottir V.G., Gudmundsdottir E., Bjornsson A., Ingvarsson B.,			
RA	Ingason A., Sigfusson S., Hardardottir H., Harvey R.P., Brunner D.,			
RA	Mutel V., Gonzalo A., Lemke G., Sainz J., Johannesson G.,			
RA	Andersson T., Gudbjartsson D., Manolescu A., Frigge M.L., Gurney M.E.,			
RA	Kong A., Gulcher J.R., Petursson H., Stefansson K.,			
RA	"Neuregulin 1 and Susceptibility to Schizophrenia.",			
RL	Am. J. Hum. Genet. 71:0-0(2002).			
CC	-1- MISCELLANEOUS: The sequence shown here is derived from an			
CC	EMBL/GenBank/DBJ third party annotation (TPA) entry.			
DR	EMBL; BK000383; DAA00041.1; -.			
DR	HSSP; Q12780; 1HRE.			
DR	GO; GO:0005102; F:receptor binding; IEA.			
DR	GO; GO:0009790; P:embryonic development; IEA.			
DR	InterPro; IPR000742; EGF 2.			
DR	InterPro; IPR006209; EGF-like.			
DR	InterPro; IPR007110; IG-like.			
DR	InterPro; IPR002154; Neuregulin.			
DR	Pfam; PF00047; igf 1.			
DR	Pfam; PF02158; Neuregulin; 1.			
DR	PRINTS; PRO1089; NEUREGULIN.			
DR	PROSITE; PS00022; EGF 1; UNKNOWN_1.			
DR	PROSITE; PS00026; EGF 3; 1.			
DR	PROSITE; PS00835; IG Like; 1.			
SQ	SEQUENCE 645 AA; 71126 MW; 04B7AF528CDD628 CRC64;			

Query Match	99.7%;	Score 3337;	DB 2;	Length 645;
Best Local Similarity	99.7%;	Pred. No. 3.3e-175;		
Matches 643;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	MSERKEGRGKGKKKRGSGKPKESAGSQSPALPPQIKEMKKSQESAGSKLVLRCTS	60	
Db	1	MSERKEGRGKGKKKRGSGKPKESAGSQSPALPPQIKEMKKSQESAGSKLVLRCTS	60	
Qy	61	SEYSLRFRKWFNGVGNELNKNKPQIKQPKGKSELKINKASLADSGEYMKVSKLGN	120	

```

Db      61 SEYSSLRFXKFKNGNELNRKNKPNQIKIQKPGKSELINKASLADSGEYMKVKISLGN 120
QY      121 DSASANITIVESNEIITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTTTGTSHLV 180
Db      121 DSASANITIVESNEIITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTTTGTSHLV 180
QY      181 KCAEKEKTEFCVANGGECFMVKDLSNPSRYLCKCPNEFTGDRCONYVMAVSPYKHLGIEFMEA 240
Db      181 KCAEKEKTEFCVANGGECFMVKDLSNPSRYLCKCPNEFTGDRCONYVMAVSPYKHLGIEFMEA 240
QY      241 BELYQKRVLTITIGICIALLVGIMCVVAYCKTKQKQKLDRLQSLRSERNMNIANG 300
Db      241 BELYQKRVLTITIGICIALLVGIMCVVAYCKTKQKQKLDRLQSLRSERNMNIANG 300
QY      301 PHHPNPPPNVQVNVQYVSKNVISSSEHIVERAEATSFSTSHYSTTAHSTTTVOTPSHSW 360
Db      301 PHHPNPPPNVQVNVQYVSKNVISSSEHIVERAEATSFSTSHYSTTAHSTTTVOTPSHSW 360
QY      361 SNGHTESILSESHSVIVMSSVENSRRHSPTGPGRLNGTGGPRCNSFLRHARETDPDSY 420
Db      361 SNGHTESILSESHSVIVMSSVENSRRHSPTGPGRLNGTGGPRCNSFLRHARETDPDSY 420
QY      421 RDSPHSERYVSAMTTPARMSPVDFHTPPSPKPPSEMSPPVSSMTVSPMAVSPFMEEE 480
Db      421 RDSPHSERYVSAMTTPARMSPVDFHTPPSPKPPSEMSPPVSSMTVSPMAVSPFMEEE 480
QY      481 RPLLLVTPPRLREKKFDHHPQPFSSFHHPNPAHDSNSLPASPLRIVEDEYEYETTOYEYPAQ 540
Db      481 RPLLLVTPPRLREKKFDHHPQPFSSFHHPNPAHDSNSLPASPLRIVEDEYEYETTOYEYPAQ 540
QY      541 EPVKKLANSRRARAKTKPNGHIANRLEVDNSTSSQSSNSESETEDESERVEDPFLGIQNPL 600
Db      541 EPVKKLANSRRARAKTKPNGHIANRLEVDNSTSSQSSNSESETEDESERVEDPFLGIQNPL 600
QY      601 AASLEATPAFLADSRTPNAGRFSTQEEIQARLSSVIANQDPIAV 645
Db      601 AASLEATPAFLADSRTPNAGRFSTQEEIQARLSSVIANQDPIAV 645

```

RESULT 2

```

QRTW3
ID      QRTW3      PRELIMINARY;      PRT;      637 AA.
AC      QRTW3;
DT      01-WAR-2004 (TREMBLrel. 26, Created)
DT      01-WAR-2004 (TREMBLrel. 26, Last sequence update)
DT      01-WAR-2004 (TREMBLrel. 26, Last annotation update)
DE      Neuregulin 1 isoform HRG-beta2.
GN      Name=NRG1;
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxId=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      PubMed=12145742;
RA      Stefansson H., Sigurdsson E., Steinthorsdottir V., Bjornsdottir S.,
RA      Ivarsson O., Ghosh S., Brynjolfsson J., Gunnarsdottir S.,
RA      Gudnadottir V.G., Gudmundsdottir E., Bjornsson A., Ingvarsson B.,
RA      Ingason A., Sigfusson S., Hardardottir H., Harvey R.P., Brunner D.,
RA      Mutel V., Gonzalo A., Lemke G., Sainz J., Johannesson G.,
RA      Andresen T., Gudbjartsson D., Manolescu A., Frigge M.L., Gurney M.E.,
RA      Kong A., Gulcher J.R., Retursson H., Stefansson K.,
RT      "Neuregulin 1 and Susceptibility to Schizophrenia.";
RL      Am. J. Hum. Genet. 71:0-0(2002).
CC      -1- MISCELLANEOUS: The sequence shown here is derived from an
CC      EMBL/GenBank/DDBJ third party annotation (rpa) entry.
DR      EMBL; BK000383; DAA0042.1; -
DR      HSP; Q12780; 1HRE
DR      GO; GO:0005102; Fireceptor binding; IEA.
DR      GO; GO:0009790; P:embryonic development; IEA.
DR      InterPro; IPR000742; EGF_2.

```

```

DR      InterPro; IPR006209; EGF like.
DR      InterPro; IPR007110; Ig-Like.
DR      InterPro; IPR002154; Neuregulin.
DR      Pfam; PF00008; EGF; 1.
DR      Pfam; PF00047; Ig; 1.
DR      Pfam; PF0158; Neuregulin; 1.
DR      PRINTS; PR01089; NEUREGULIN.
DR      PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR      PROSITE; PS00026; EGF_3; 1.
DR      PROSITE; PS00835; IG_LIKE; 1.
DR      SEQUENCE 637 AA; 70169 MW; 0B89DF6E2ED70333 CRC64;

Query Match      98.0%; Score 3280; DB 2; Length 637;
Best Local Similarity 98.4%; Pred. No. 4.4e-172;
Matches 635; Conservative 1; Mismatches 1; Indels 8; Gaps 1;

QY      1 MSERKEGKGKGGKKGKSGKSPESAAGSQSPALPOLKEMKQESAAAGSKLVLCRTS 60
Db      1 MSERKEGKGKGGKKGKSGKSPESAAGSQSPALPOLKEMKQESAAAGSKLVLCRTS 60
QY      61 SEYSSLRFXKFKNGNELNRKNKPNQIKIQKPGKSELINKASLADSGEYMKVKISLGN 120
Db      61 SEYSSLRFXKFKNGNELNRKNKPNQIKIQKPGKSELINKASLADSGEYMKVKISLGN 120
QY      121 DSASANITIVESNEIITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTTTGTSHLV 180
Db      121 DSASANITIVESNEIITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTTTGTSHLV 180
QY      181 KCAEKEKTEFCVANGGECFMVKDLSNPSRYLCKCPNEFTGDRCONYVMAVSPYKHLGIEFMEA 240
Db      181 KCAEKEKTEFCVANGGECFMVKDLSNPSRYLCKCPNEFTGDRCONYVMAVSPYKHLGIEFMEA 240
QY      241 BELYQKRVLTITIGICIALLVGIMCVVAYCKTKQKQKLDRLQSLRSERNMNIANG 300
Db      241 BELYQKRVLTITIGICIALLVGIMCVVAYCKTKQKQKLDRLQSLRSERNMNIANG 300
QY      301 PHHPNPPPNVQVNVQYVSKNVISSSEHIVERAEATSFSTSHYSTTAHSTTTVOTPSHSW 360
Db      301 PHHPNPPPNVQVNVQYVSKNVISSSEHIVERAEATSFSTSHYSTTAHSTTTVOTPSHSW 360
QY      361 SNGHTESILSESHSVIVMSSVENSRRHSPTGPGRLNGTGGPRCNSFLRHARETDPDSY 420
Db      361 SNGHTESILSESHSVIVMSSVENSRRHSPTGPGRLNGTGGPRCNSFLRHARETDPDSY 420
QY      421 RDSPHSERYVSAMTTPARMSPVDFHTPPSPKPPSEMSPPVSSMTVSPMAVSPFMEEE 480
Db      421 RDSPHSERYVSAMTTPARMSPVDFHTPPSPKPPSEMSPPVSSMTVSPMAVSPFMEEE 480
QY      481 RPLLLVTPPRLREKKFDHHPQPFSSFHHPNPAHDSNSLPASPLRIVEDEYEYETTOYEYPAQ 540
Db      481 RPLLLVTPPRLREKKFDHHPQPFSSFHHPNPAHDSNSLPASPLRIVEDEYEYETTOYEYPAQ 540
QY      541 EPVKKLANSRRARAKTKPNGHIANRLEVDNSTSSQSSNSESETEDESERVEDPFLGIQNPL 600
Db      541 EPVKKLANSRRARAKTKPNGHIANRLEVDNSTSSQSSNSESETEDESERVEDPFLGIQNPL 600
QY      601 AASLEATPAFLADSRTPNAGRFSTQEEIQARLSSVIANQDPIAV 645
Db      593 AASLEATPAFLADSRTPNAGRFSTQEEIQARLSSVIANQDPIAV 637

```

RESULT 3

```

NRG1_HUMAN
ID      NRG1_HUMAN      STANDARD;      PRT;      639 AA.
AC      Q02297; Q14667; P98202; Q02298; Q02299; Q07110; Q07111; Q12779;
AC      Q12780; Q12781; Q12782; Q12783; Q12784; Q9UPE3;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      25-OCT-2004 (Rel. 45, Last annotation update)
DE      Pre-neuregulin-1 precursor (Pro-NRG1) [contains: Neuregulin-1 (Neu
DE      differentiation factor) (Herregulin) (HRG) (breast cancer cell
DE      (ARIA) (Sensory and motor neuron-derived factor) (Glial growth

```

DE factor]]].
GN Name=NRGI; Synonyms=GGF, HGL, HRGA, NDF, SMDF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORMS 1; 6; 7 AND 8), AND PARTIAL SEQUENCE.
RP MEDLINE=92271253; PubMed=1350381;
RA Holmes W.E., Sliwkowski M.X., Akita R.W., Henzel W.J., Lee J.,
RA Park J.W., Yansura D., Abadi N., Raab H., Lewis G.D., Shepard H.M.,
RA Kuang W.-J., Wood W.I., Goeddel D.V., Vanden R.L.;
RT "Identification of heregulin, a specific activator of p185erbB2.";
RL Science 256:1205-1210(1992).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORMS 2; 3; 4; 6; 7 AND 8).
RP TISSUE=Kidney adenocarcinoma, and Pituitary;
RX MEDLINE=94158863; PubMed=7509448;
RA Wen D., Suggs S.V., Karunakaran D., Liu N., Cupples R.L., Luo Y.,
RA Janssen A.M., Ben-Baruch N., Tollinger D.B., Jacobsen V.L.,
RA Meng S.-Y., Lu H.S., Hu S., Chang D., Yang W., Yanigahara D.,
RA Koski R.A., Yarden Y.;
RT "Structural and functional aspects of the multiplicity of Neu
RT differentiation factors.";
RL Mol. Cell. Biol. 14:1909-1919(1994).
RN [3]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RP MEDLINE=92208945; PubMed=1348215; DOI=10.1016/0092-8674(92)90131-U;
RA Peles E., Bacus S.S., Koski R.A., Lu H.S., Wen D., Ogden S.G.,
RA Levy R.B., Yarden Y.;
RT "Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein
RT that induces differentiation of mammary tumor cells.";
RL Cell 69:205-216(1992).
RN [4]
RN SEQUENCE FROM N.A. (ISOFORMS 8 AND 9).
RP TISSUE=Brain;
RX MEDLINE=93205115; PubMed=8096067; DOI=10.1038/362312a0;
RA Marchionni M.A., Goodearl A.D.J., Chen M.-S., Birmingham-McDonogh O.,
RA Kirk C., Hendricks M., Danehy F., Misumi D., Sudhalter J.,
RA Kobayashi K., Wroblewski D., Lynch C., Baldaasare M., Hiles I.,
RA Davis J.B., Hsuan J.-J., Totty N.P., Oteu M., McBurney R.N.,
RA Waterfield M.D., Stroobant P., Gwynne D.;
RT "Glial growth factors are alternatively spliced erbB2 ligands
RT expressed in the nervous system.";
RL Nature 362:312-318(1993).
RN [5]
RN SEQUENCE FROM N.A. OF GAMMA-HERGULIN FUSION PROTEIN.
RP TISSUE=Breast cancer;
RX MEDLINE=97472144; PubMed=9333014; DOI=10.1038/sj.onc.1201317;
RA Schaefer G., Fitzpatrick V.D., Sliwkowski M.X.;
RT "Gamma-hergulin: a novel heregulin isoform that is an autocrine
RT growth factor for the human breast cancer cell line, MDA-MB-175.";
RL Oncogene 15:1385-1394(1997).
RN [6]
RN SEQUENCE OF 1-210 FROM N.A.
RP Schoumacher F., Herzer S., Flury N., Kueng W., Mueller H.,
RA Eppenberger U.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [7]
RN SEQUENCE OF 19-27.
RP MEDLINE=93366731; PubMed=7689552;
RA Culouscou J.-M., Plozman G.D., Carlton G.W., Green J.M., Shoyab M.;
RT "Characterization of a breast cancer cell differentiation factor that
RT specifically activates the HER4/p180erbB4 receptor.";
RL J. Biol. Chem. 268:18407-18410(1993).
RN [8]
RN CHROMOSOMAL TRANSLOCATION.
RX MEDLINE=99455251; PubMed=10523851; DOI=10.1038/sj.onc.1202950;
RA Wang X.-Z., Jolicoeur E.M., Conte N., Chaffanet M., Zhang Y.,
RA Mozziconacci M.-J., Feiner H., Birnbaum D., Febusque M.-J., Ron D.;
RT "Gamma-hergulin is the product of a chromosomal translocation fusing
RT the DOC4 and HGL/NRG1 genes in the MDA-MB-175 breast cancer cell
RT line.";

OncoGene 18:5718-5721(1999).
RN [9]
RN CHROMOSOMAL TRANSLOCATION.
RX MEDLINE=20065180; PubMed=10597312; DOI=10.1038/sj.onc.1203136;
RA Liu X., Baker E., Eyre H.J., Sutherland G.R., Zhou M.;
RT "Gamma-hergulin: a fusion gene of DOC-4 and neuregulin-1 derived from
RT a chromosome translocation.";
RL Oncogene 18:7110-7114(1999).
RN [10]
RN STRUCTURE BY NMR OF 175-241 (ISOFORM 1).
RX MEDLINE=94341264; PubMed=8062828;
RA Nagata K., Kohda D., Hatanaka H., Ichikawa S., Matsuda S.,
RA Yamamoto T., Suzuki A., Inagaki F.;
RT "Solution structure of the epidermal growth factor-like domain of
RT heregulin-alpha, a ligand for p180erbB-4.";
RL EMBO J. 13:3517-3523(1994).
CC -1- FUNCTION: Direct ligand for ERBB3 and ERBB4 tyrosine kinase
CC receptors. Concomitantly recruits ERBB1 and ERBB2 coreceptors,
CC resulting in ligand-stimulated tyrosine phosphorylation and
CC activation of the ERBB receptors. The multiple isoforms perform
CC diverse functions such as inducing growth and differentiation of
CC epithelial, glial, neuronal, and skeletal muscle cells; inducing
CC expression of acetylcholine receptor in synaptic vesicles during
CC the formation of the neuromuscular junction; stimulating
CC lobuloalveolar budding and milk production in the mammary gland
CC and inducing differentiation of mammary tumor cells; stimulating
CC Schwann cell proliferation; implication in the development of the
CC myocardium such as trabeculation of the developing heart.
CC -1- SUBUNIT: The cytoplasmic domain interacts with the LIM domain
CC region of LIMK1 (by similarity).
CC -1- SUBCELLULAR LOCATION: Exists as an type I membrane protein and as
CC a proteolytically released soluble growth factor form. The
CC membrane-bound form does not seem to be active. The secreted
CC isoform 9 has a signal peptide. The isoform 8 may be nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=9;
CC Comment=Additional isoforms seem to exist. Isoforms have been
CC classified as type I NRGs (isoforms with an Ig domain and a
CC glycosylation domain, isoforms 1-8), type II NRGs (isoforms with
CC an Ig domain but no glycosylation domain, isoform 9) and type
CC III NRGs (isoforms with a Cys-rich domain, isoform 10). All
CC these isoforms perform distinct tissue-specific functions;
CC Name=1; Synonyms=Alpha;
CC IsoId=Q02297-1; Sequence=Displayed;
CC Name=2; Synonyms=Alpha1A;
CC IsoId=Q02297-2; Sequence=VSP_003431;
CC Name=3; Synonyms=Alpha2B;
CC IsoId=Q02297-3; Sequence=VSP_003434, VSP_003435;
CC Name=4; Synonyms=Alpha3;
CC IsoId=Q02297-4; Sequence=VSP_003432, VSP_003433;
CC Name=6; Synonyms=Betal, BetalA;
CC IsoId=Q02297-6; Sequence=VSP_003428;
CC Name=7; Synonyms=Beta2;
CC IsoId=Q02297-7; Sequence=VSP_003427;
CC Name=8; Synonyms=Beta3, GGFHPB1;
CC IsoId=Q02297-8; Sequence=VSP_003429, VSP_003430;
CC Name=9; Synonyms=GGF2, GGFHPB2;
CC IsoId=Q02297-9; Sequence=VSP_003425, VSP_003426, VSP_003429,
CC VSP_003430;
CC Name=10; Synonyms=SMDF;
CC IsoId=Q15491-1; Sequence=External;
CC -1- TISSUE SPECIFICITY: Type I isoforms are the predominant forms
CC expressed in the endocardium. Isoform alpha is expressed in
CC breast, ovary, testis, prostate, heart, skeletal muscle, lung,
CC placenta liver, kidney, salivary gland, small intestine and brain,
CC but not in uterus, stomach, pancreas, and spleen. Isoform 3 is the
CC predominant form in mesenchymal cells and in nonneuronal organs,
CC whereas isoform 5 is the major neuronal form. Isoform 8 is
CC expressed in spinal cord and brain. Isoform 9 is the major form in
CC skeletal muscle cells; in the nervous system it is expressed in
CC spinal cord and brain. Also detected in adult heart, placenta,
CC lung, liver, kidney, and pancreas.
CC -1- DEVELOPMENTAL STAGE: Detectable at early embryonic ages.


```
Db 1 MSERKEGRGKGKGGKKGKGPESAAQSQSPALPPRLKEMKMSQSAAGSKVLRLCETS 60
QY 61 SEYSSLRFKWFKNGNELNRKQPNKIQKKPKGKSELRLINKASLADSGEYMKVSKLGN 120
Db 61 SEYSSLRFKWFKNGNELNRKQPNKIQKKPKGKSELRLINKASLADSGEYMKVSKLGN 120
QY 121 DSASANITIVESNEIITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTGTSHLV 180
Db 121 DSASANITIVESNEIITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTGTSHLV 180
QY 181 KCAEKEKTCVNGGECFVKDLNPSRYLCKPNEFTGDRCONYMASFYKHLGIEFMEA 240
Db 181 KCAEKEKTCVNGGECFVKDLNPSRYLCKPNEFTGDRCONYMASFYKHLGIEFMEA 240
QY 241 EELYQKRVLTITIGICIALLVGIMCVVAYCTKQKPKGKSELRLINKASLADSGEYMKVSKLGN 300
Db 236 EELYQKRVLTITIGICIALLVGIMCVVAYCTKQKPKGKSELRLINKASLADSGEYMKVSKLGN 295
QY 301 PPHPNPPPPENVOLVQVSKNVISSSEHIVERAEATSFSTSHYTSTAHHSTTTVTQTPSHW 360
Db 296 PPHPNPPPPENVOLVQVSKNVISSSEHIVERAEATSFSTSHYTSTAHHSTTTVTQTPSHW 355
QY 361 SNGHTESILSESHSVIMSSVENSRHSSPTGGRGLNGTGPPRECNSFLRHARETPDSY 420
Db 356 SNGHTESILSESHSVIMSSVENSRHSSPTGGRGLNGTGPPRECNSFLRHARETPDSY 415
QY 421 RDSPHSERVVSAMTTPARMSPVDFHTPPSPKSPPESEMPVSSMTVMPMSVAPPMEE 480
Db 416 RDSPHSERVVSAMTTPARMSPVDFHTPPSPKSPPESEMPVSSMTVMPMSVAPPMEE 475
QY 481 RPLLLVTPRLREKKFDHHPQPFSSFHHPNPAHDSNLSPLASPLRIVEDEEYETTQYEPAQ 540
Db 476 RPLLLVTPRLREKKFDHHPQPFSSFHHPNPAHDSNLSPLASPLRIVEDEEYETTQYEPAQ 535
QY 541 EPVKKLANSRRARCTKPNGHIANRLEVDNNTSSQSSNSETEDEVRGDTTFLGIQNPL 600
Db 536 EPVKKLANSRRARCTKPNGHIANRLEVDNNTSSQSSNSETEDEVRGDTTFLGIQNPL 595
QY 601 AASLEATPAFLADSRTPAGRSTQEEIQAARLSSVIANQDPIAV 645
Db 596 AASLEATPAFLADSRTPAGRSTQEEIQAARLSSVIANQDPIAV 640
```

RESULT 5

```
Q6DR98
ID Q6DR98 PRELIMINARY; PRT; 645 AA.
AC Q6DR98;
DT 25-OCT-2004 (T:EMBLrel. 28, Created)
DT 25-OCT-2004 (T:EMBLrel. 28, Last sequence update)
DE Neuregulin-1 type I betal-a.
GN Name:Nrgl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57Bl6J;
RA Anton E.S., Ghassghaei H.T., Weber J.L., McCann C., Fischer T.M.,
RA Cheung I.D., Gassmann M., Messing A., Klein R., Schwab M.H.,
RA Lloyd K.C., Lai C.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
EMBL: AY648976; AAT68241.1;
DR GO; GO:0005102; F:receptor binding; IEA.
DR GO; GO:0009790; P:embryonic development; IEA.
DR InterPro; IPR000742; EGF-2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
```

```
DR InterPro; IPR002154; Neuregulin.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF02158; Neuregulin; 1.
DR PRINTS; PR01089; NEUREGULIN.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS00022; EGF 1; UNKNOWN_1.
DR PROSITE; PS00026; EGF 3; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR EGF-like domain.
SQ SEQUENCE 645 AA; 71381 MW; 75755AEF73F55047 CRC64;

Query Match 94.1%; Score 3149; DB 2; Length 645;
Best Local Similarity 93.0%; Pred. No. 6.9e-165;
Matches 600; Conservative 21; Mismatches 24; Indels 0; Gaps 0;

QY 1 MSERKEGRGKGKGGKKGKGPESAAQSQSPALPPRLKEMKMSQSAAGSKVLRLCETS 60
Db 1 MSERKEGRGKGKGGKKGKGPESAAQSQSPALPPRLKEMKMSQSAAGSKVLRLCETS 60
QY 61 SEYSSLRFKWFKNGNELNRKQPNKIQKKPKGKSELRLINKASLADSGEYMKVSKLGN 120
Db 61 SEYSSLRFKWFKNGNELNRKQPNKIQKKPKGKSELRLINKASLADSGEYMKVSKLGN 120
QY 121 DSASANITIVESNEIITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTGTSHLV 180
Db 121 DSASANITIVESNEIITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTGTSHLV 180
QY 181 KCAEKEKTCVNGGECFVKDLNPSRYLCKPNEFTGDRCONYMASFYKHLGIEFMEA 240
Db 181 KCAEKEKTCVNGGECFVKDLNPSRYLCKPNEFTGDRCONYMASFYKHLGIEFMEA 240
QY 241 EELYQKRVLTITIGICIALLVGIMCVVAYCTKQKPKGKSELRLINKASLADSGEYMKVSKLGN 300
Db 241 EELYQKRVLTITIGICIALLVGIMCVVAYCTKQKPKGKSELRLINKASLADSGEYMKVSKLGN 300
QY 301 PPHPNPPPPENVOLVQVSKNVISSSEHIVERAEATSFSTSHYTSTAHHSTTTVTQTPSHW 360
Db 301 PPHPNPPPPENVOLVQVSKNVISSSEHIVERAEATSFSTSHYTSTAHHSTTTVTQTPSHW 360
QY 361 SNGHTESILSESHSVIMSSVENSRHSSPTGGRGLNGTGPPRECNSFLRHARETPDSY 420
Db 361 SNGHTESILSESHSVIMSSVENSRHSSPTGGRGLNGTGPPRECNSFLRHARETPDSY 420
QY 421 RDSPHSERVVSAMTTPARMSPVDFHTPPSPKSPPESEMPVSSMTVMPMSVAPPMEE 480
Db 421 RDSPHSERVVSAMTTPARMSPVDFHTPPSPKSPPESEMPVSSMTVMPMSVAPPMEE 480
QY 481 RPLLLVTPRLREKKFDHHPQPFSSFHHPNPAHDSNLSPLASPLRIVEDEEYETTQYEPAQ 540
Db 481 RPLLLVTPRLREKKFDHHPQPFSSFHHPNPAHDSNLSPLASPLRIVEDEEYETTQYEPAQ 540
QY 541 EPVKKLANSRRARCTKPNGHIANRLEVDNNTSSQSSNSETEDEVRGDTTFLGIQNPL 600
Db 541 EPVKKLANSRRARCTKPNGHIANRLEVDNNTSSQSSNSETEDEVRGDTTFLGIQNPL 600
QY 601 AASLEATPAFLADSRTPAGRSTQEEIQAARLSSVIANQDPIAV 645
Db 601 AASLEATPAFLADSRTPAGRSTQEEIQAARLSSVIANQDPIAV 645
```

RESULT 6

```
NRG1_RAT
ID NRG1_RAT STANDARD; PRT; 662 AA.
AC P43323; P43324; P43325; P43326; P43327; P43328;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Pro-neuregulin-1 precursor (Pro-NRG1) [Contains: Neuregulin-1 (Neu
DE differentiation factor) (Heregulin) (HRG) (Acetylcholine receptor
DE inducing activity) (ARIA) (Sensory and motor neuron-derived factor)
```

DE (Glial growth factor)]].
GN Name=Nrg1; Synonyms=Ndf;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX TISSUE=Fibroblast;
RA MEDLINE=9415863; PubMed=7509448;
RA Wen D., Suggs S.V., Karunakaran D., Liu N., Cupples R.L., Luo Y.,
RA Janssen A.M., Ben-Baruch N., Trollinger D.B., Jacobsen V.L.,
RA Meng S.-Y., Lu H.S., Hu S., Chang D., Yang W., Yanigahara D.,
RA Koski R.A., Yarden Y.;
RT "Structural and functional aspects of the multiplicity of Neu
RT differentiation factors";
RL Mol. Cell. Biol. 14:1909-1919(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA2C/NDP44), AND PARTIAL SEQUENCE.
RX TISSUE=Fibroblast;
RA MEDLINE=92208945; PubMed=1349853; DOI=10.1016/0092-8674(92)90456-M;
RA Wen D., Peles E., Cupples R., Suggs S.V., Bacus S.S., Luo Y.,
RA Trull G., Hu S., Silbiger S.M., Levy R.B., Koski R.A., Lu H.S.,
RA Yarden Y.;
RT "Neu differentiation factor: a transmembrane glycoprotein containing
RT an EGF domain and an immunoglobulin homology unit";
RL Cell 69:559-572(1992).
RN [3]
RP SEQUENCE OF 14-36.
RX MEDLINE=92208945; PubMed=1349853; DOI=10.1016/0092-8674(92)90456-M;
RA Peles E., Bacus S.S., Koski R.A., Lu H.S., Wen D., Ogden S.G.,
RA Levy R.B., Yarden Y.;
RT "Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein
RT that induces differentiation of mammary tumor cells";
RL Cell 69:205-216(1992).
RN [4]
RP REGULATION OF PROCESSING (ISOFORM ALPHA2C/NDP44).
RX MEDLINE=99069430; PubMed=9852099; DOI=10.1074/jbc.273.51.34335;
RA Liu X., Hwang H., Cao L., Wen D., Liu N., Graham R.M., Zhou M.;
RT "Release of the neu/HER-2 stimulatory ligand requires its
RT cytoplasmic tail";
RL J. Biol. Chem. 273:34335-34340(1998).
RN [5]
RP INTERACTION WITH LIMK1.
RX MEDLINE=98352096; PubMed=9685409; DOI=10.1074/jbc.273.32.20525;
RA Wang J.Y., Frenzel K.E., Wen D., Falls D.L.;
RT "Transmembrane neu/HER-2 interact with LIM kinase 1, a cytoplasmic
RT protein kinase implicated in development of visuospatial cognition";
RL J. Biol. Chem. 273:20525-20534(1998).
CC -!- FUNCTION: Direct ligand for ERBB3 and ERBB4 tyrosine kinase
CC receptors. Concomitantly recruits ERBB1 and ERBB2 coreceptors,
CC resulting in ligand-stimulated tyrosine phosphorylation and
CC activation of the ERBB receptors. The multiple isoforms perform
CC diverse functions such as inducing growth and differentiation of
CC epithelial, glial, neuronal, and skeletal muscle cells; inducing
CC expression of acetylcholine receptor in synaptic vesicles during
CC the formation of the neuromuscular junction; stimulating
CC lobuloalveolar budding and milk production in the mammary gland
CC and inducing differentiation of mammary tumor cells; stimulating
CC Schwann cell proliferation; implication in the development of the
CC myocardium such as trabeculation of the developing heart (By
CC similarity).
CC -!- SUBUNIT: The cytoplasmic domain interacts with the LIM domain
CC region of LIMK1.
CC -!- SUBCELLULAR LOCATION: Exists as a type I membrane protein and as a
CC proteolytically released soluble growth factor form. The membrane-
CC bound form does not seem to be active.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=8;
CC Comment=Additional isoforms seem to exist;
CC Name=Beta4; Synonyms=NDP42A;
CC IsoId=P43322-1; Sequence=Displayed;
CC Name=Alpha2A; Synonyms=NDP38;

CC IsoId=P43322-2; Sequence=VSP_003436;
CC Name=Alpha2B; Synonyms=NDP19;
CC IsoId=P43322-3; Sequence=VSP_003436, VSP_003443, VSP_003444;
CC Name=Alpha2C; Synonyms=NDP44;
CC IsoId=P43322-4; Sequence=VSP_003436, VSP_003442;
CC Name=Beta1;
CC IsoId=P43322-5; Sequence=VSP_003437;
CC Name=Beta2; Synonyms=NDP40;
CC IsoId=P43322-6; Sequence=VSP_003440, VSP_003441;
CC Name=Beta2A; Synonyms=NDP22;
CC IsoId=P43322-7; Sequence=VSP_003440;
CC Name=Beta3; Synonyms=NDP4;
CC IsoId=P43322-8; Sequence=VSP_003438, VSP_003439;
CC -!- TISSUE SPECIFICITY: Widely expressed. Most tissues contain alpha2A
CC and alpha2B isoforms. Alpha2 and beta2 are the predominant forms
CC in mesenchymal and nonneural organs. Beta1 is enriched in brain
CC and spinal cord, but not in muscle and heart. Alpha2C is highly
CC expressed in spinal cord, moderately in lung, brain, ovary, and
CC stomach, in low amounts in the kidney, skin and heart and not
CC detected in the liver, spleen, and placenta.
CC -!- DOMAIN: The cytoplasmic domain may be involved in the regulation of
CC trafficking and proteolytic processing. Regulation of the
CC proteolytic processing involves initial intracellular domain
CC dimerization.
CC -!- DOMAIN: ERBB receptor binding is elicited entirely by the EGF-like
CC domain.
CC -!- PTM: Proteolytic cleavage close to the plasma membrane on the
CC external face leads to the release of the soluble growth factor
CC form.
CC -!- PTM: Extensive glycosylation precedes the proteolytic cleavage.
CC -!- SIMILARITY: Belongs to the neu/HER-2 family.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U02315; AAA19940.1; -
CC EMBL; U02316; AAA19941.1; -
CC EMBL; U02317; AAA19942.1; -
CC EMBL; U02318; AAA19943.1; -
CC EMBL; U02319; AAA19944.1; -
CC EMBL; U02320; AAA19945.1; -
CC EMBL; U02321; AAA19946.1; -
CC EMBL; U02322; AAA19947.1; -
CC EMBL; U02323; AAA19948.1; -
CC EMBL; U02324; AAA19949.1; -
CC EMBL; M92430; -; NOT ANNOTATED_CDS.
CC PIR; I61718; I61718.
CC PIR; I61719; I61719.
CC PIR; I61722; I61722.
CC HSSP; Q12780; 1HRE.
CC RGD; 621341; Nrg1.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003598; IG_C2.
CC InterPro; IPR002154; Neuregulin.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF02158; Neuregulin; 1.
CC PRINTS; PR01089; NEUREGULIN.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00408; IGC2; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; FALSE_NEG.
CC PROSITE; PS50026; EGF_3; 1.

Best Local Similarity 89.8%; Pred. No. 1e-156;
Matches 579; Conservative 23; Mismatches 23; Indels 20; Gaps 2;

QY 20 SKKESAGSOSP-----ALPPQKEMKSOESAGSKVLRCETS 60
DB 139 SGRAFPAPFASPPFLEGTGNLKKEVSRVLCRKCALPPRLKEMKSOESAGSKVLRCETS 198
QY 61 SEYSSIRFKWFGNGELNKRKNPQIKIOKPKGSELRLINKASLADSGEYMKCVISKLN 120
DB 199 SEYSSIRFKWFGNGELNKRKNPQIKIOKPKGSELRLINKASLADSGEYMKCVISKLN 258
QY 121 DSANANITIVESNEITGMPASTEGAYVSSSPIRISVSTEGANSSSTSTTTCTSHLV 180
DB 259 DSANANITIVESNEITGMPASTEGAYVSSSPIRISVSTEGANSSSTSTTTCTSHLI 318
QY 181 KCAEKEKTCVNGGECFVKDLSNPSRYLCKCPNEFTGDRCONYMASFYKHLGIEFMEA 240
DB 319 KCAEKEKTCVNGGECFVKDLSNPSRYLCKCPNEFTGDRCONYMASFYKHLGIEFMEA 378
QY 241 BELYQKRVLTITIGICALLVWGMVAYCKTKKQKLDRLRSLRSERNMNIANG 300
DB 379 BELYQKRVLTITIGICALLVWGMVAYCKTKKQKLDRLRSLRSERNMNIANG 438
QY 301 PHHPNPPENVOLVQYVSKNVISSEHIVEREAEATSFSTSHYTSTAHSHTTQTTPSHSW 360
DB 439 PHHPNPPENVOLVQYVSKNVISSEHIVEREAEATSFSTSHYTSTAHSHTTQTTPSHSW 498
QY 361 SNGHTESILSESHSVIVMSVENSRRHSPTGPRGLNGTGPSCNCSFLRHARETPDSY 420
DB 499 SNGHTESILSESHSVIVMSVENSRRHSPTGPRGLNGTGPSCNCSFLRHARETPDSY 557
QY 421 RDSPHSERVSVANTTPARMSPVDFHTPSPKPPSEMPPVSSMTVSPMSMAVSFMEEE 480
DB 558 RDSPHSERVSVANTTPARMSPVDFHTPSPKPPSEMPPVSSMTVSPMSMAVSFMEEE 617
QY 481 RPLLIVTPRLREKXPDHHPQOFSSFHNPANHSNLSLSPASPLRIVEDEEYETQYEPQAQ 540
DB 618 RPLLIVTPRLREKXPDHHPQOFSSFHNPANHSNLSLSPASPLRIVEDEEYETQYEPQAQ 677
QY 541 EVVKLANSRRAKRPKNGHIANRLEVDNSTSSQSSNESEFEDERVDGDPFLGIQNP 600
DB 678 EVVKVNTVRRRAKRPKNGHIANRLEVDNSTSSQSSNESEFEDERVDGDPFLGIQNP 737
QY 601 AASLEATPAFLADSRTPAGRFSQEEIQARLSSLVIANQDPIAV 645
DB 738 AASLEVAPAFRLAERTNPAGRFSQEEIQARLSSLVIANQDPIAV 782

RESULT 8
NRG1 CHICK STANDARD; PRT; 602 AA.
AC Q05199; 073750; 073751; 073752;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Pro-neuregulin-1 precursor (Pro-NRG1) [Contains: Neuregulin-1
DE (Acetylcholine receptor inducing activity) (ARIA)].
GN Name=NRG1; Synonyms=ARIA;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RC STRAIN=White Leghorn; TISSUE=Brain;
RX MEDLINE=93201602; PubMed=8453670; DOI=10.1016/0092-8674(93)90407-H;
RA Falls D.L., Rosen K.M., Corfas G., Lane W.S., Fischbach G.D.;
RT "ARIA, a protein that stimulates acetylcholine receptor synthesis, is
RL a member of the neu ligand family.";
RN Cell 72:801-815(1993).
RP SEQUENCE FROM N.A. (ISOFORMS 2; 3 AND 4).

RC TISSUE=Brain, and Spinal cord;
RX MEDLINE=98150951; PubMed=9491987; DOI=10.1016/S0896-6273(00)80454-7;
RA Yang X., Kuo Y., Devay P., Yu C., Role L.;
RT "A cysteine-rich isoform of neuregulin controls the level of
RT expression of neuronal nicotinic receptor channels during
RL synaptogenesis.";
RL Neuron 20:255-270(1998).
CC -!- FUNCTION: Direct ligand for the ERBB tyrosine kinase receptors.
CC The multiple isoforms perform diverse functions: cysteine-rich
CC domain containing isoforms (isoforms 2-4) probably regulate the
CC expression of nicotinic acetylcholine receptors at developing
CC interneuronal synapses. The Ig-NRG isoform is required for the
CC initial induction and/or maintenance of the mature levels of
CC acetylcholine receptors at neuromuscular synapses.
CC -!- SUBCELLULAR LOCATION: Exists as a type I membrane protein and as a
CC proteolytically released soluble growth factor form. The membrane-
CC bound form does not seem to be active (by similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Additional isoforms seem to exist;
CC Name=1; Synonyms=ARIA, IG-NRG;
CC IsoId=Q05199-1; Sequence=Displayed;
CC Note=Contains an Ig-like domain;
CC Name=2; Synonyms=CRD-NRG-BETA1A;
CC IsoId=Q05199-2; Sequence=VSP_003445;
CC Note=The EGF-like domain is replaced by a cysteine-rich domain
CC (CRD);
CC Name=3; Synonyms=CRD-NRG-BETA2A;
CC IsoId=Q05199-3; Sequence=VSP_003445, VSP_003446;
CC Note=The EGF-like domain is replaced by a cysteine-rich domain
CC (CRD);
CC Name=4; Synonyms=CRD-NRG-BETA2B;
CC IsoId=Q05199-4; Sequence=VSP_003445, VSP_003446, VSP_003447,
CC VSP_003448;
CC Note=The EGF-like domain is replaced by a cysteine-rich domain
CC (CRD);
CC -!- DEVELOPMENTAL STAGE: Isoforms 2-4 are detected at embryonic day 4
CC (ED4) in both visceral and somatic motor neurons of spinal cord
CC and is highest at ED6. Isoform 1 is not expressed until ED 6 in
CC spinal cord. At ED 11 both isoforms display comparable levels.
CC -!- DOMAIN: The cytoplasmic domain may be involved in the regulation
CC of trafficking and proteolytic processing. Regulation of the
CC proteolytic processing involves initial intracellular domain
CC dimerization (by similarity).
CC -!- DOMAIN: ERBB receptor binding is elicited entirely by the EGF-like
CC domain.
CC -!- PTM: Proteolytic cleavage close to the plasma membrane on the
CC external face leads to the release of the soluble growth factor
CC form.
CC -!- PTM: Extensive glycosylation precedes the proteolytic cleavage (by
CC similarity).
CC -!- SIMILARITY: Belongs to the neuregulin family.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/ebis/announce/>
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; L11264; AAA49037.1; --
CC EMBL; AF045654; AAC05670.1; --
CC EMBL; AF045655; AAC05671.1; --
CC EMBL; AF045656; AAC05672.1; --
CC PIR; A45769; A45769.
CC HSP; Q12780; 1HRE.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR007110; IG-like.
CC

```

DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR002154; Neuregulin.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF02158; Neuregulin; 1.
DR PRINTS; PR01089; NEUREGULIN.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Alternative splicing; Direct protein sequencing; EGF-like domain;
KW Glycoprotein; Growth factor; Immunoglobulin domain; Transmembrane.
FT CHAIN 1 602
FT CHAIN 1 205
FT DOMAIN 1 206
FT TRANSMEM 207 229
FT DOMAIN 230 602
FT DOMAIN 29 123
FT DOMAIN 125 136
FT DOMAIN 137 181
FT DISULFID 49 105
FT DISULFID 141 155
FT DISULFID 149 169
FT DISULFID 171 180
FT CARBOHYD 21 21
FT CARBOHYD 113 113
FT CARBOHYD 126 126
FT VARSPLIC 1 127
FT VARSPLIC 1 127
FT VARSPLIC 191 198
FT VARSPLIC 388 405
FT VARSPLIC 406 602
FT SEQUENCE 602 AA; 67453 MW; 4183C0E56CE5D346 CRC64;

Query Match 75.5%; Score 2526.5; DB 1; Length 602;
Best Local Similarity 77.2%; Pred. No. 9.7e-131;
Matches 480; Conservative 54; Mismatches 51; Indels 37; Gaps 4;

QY 25 ESAGSQSPALPOLKEMKSOBSAAGSKLVLCETSSSEYSSLRPFKFKNGNELNKNKQPO 84
DB 17 QTDVNSYNTVPKLEKMKQNEVAVGQKLVLCETTSSEYPAFLRFKWLKNGKEITKNRPE 76
QY 85 NIKIQKPGK-BELRLINKASLADSGYMKCVKSLGNDASANITIVESNEIITGMPAST 143
DB 77 NVKIPKQKSELHYRLATLADAGEYACRVSKLGNDSKASVIITDN----- 126
QY 144 EGAYVSSSEPIRISVSTEGANTSSSTSTSTGTHLVKCAEKEKTCFVNGGECFVWKDLS 203
DB 127 -----ATSTSTGTHLTKCDIKQKAFVNGGECFVWKDLP 162
QY 204 NPSRYLCKPNBFTGDRCONYMASFYKHLGIEFMEAEELYQKRVLTITGICALLVVGII 263
DB 163 NPPLYLRCPNBFTGDRCONYMASFYKHLGIEFMEAEELYQKRVLTITGICALLVVGII 222
QY 264 MCVVAYCKTKOKKHLDRSLRSERNNMNIANGPHHPNPPNPNVQVYVSKNVI 323
DB 223 MCVVAYCKTKOKKHLDRSLRSERNNMNIANGPHHPNPPNPNVQVYVSKNII 282
QY 324 SSEHIVERAEATSFSTSHYTTTQTPSHSWNGHTESTILSESHSVLVSSSVEN 383

```

```

283 SSERVVERETETSFSTSHYTTTTHSMVTQTTPSHSWNGHTESTILSESHSVLVSSSVEN 342
384 SRHSSGTPGPRGLNGTGPPECNCFRLRHARETTPDSYRDSPHSERVVSAMTTTPARMSVD 443
343 SRHTSPT-GPRGLNGIGGPRGNSFLRHARETTPDSYRDSPHSERVVSAMTTTPARMSVD 401
444 FHTSPSPKSPSEMSPPVSSMTVSPMAVSPFMEERPLLLVTPPRLEKKFDDHHPQOF 503
402 FHTTTPSPKSPSEMSPPVSSLTISPSVAVSPFMDERPLLLVTPPRLE-KYDNLQOF 460
504 SFHNPANDNSLSPSLRIVEDEYETQYEPAPQPVKLANSSRAKRTKPNGHIAN 563
461 NSFHNPTHEGNSLPPSLRIVEDEYETQYEPAPQPVKLTNSRRVKRTKPNGHISS 520
564 RLEVDNSTSSOSSSESTEDERVCEDTPELGIQNPPLAASLEATPAFLADSRTPAGRF 623
521 RVEDSDTSSQSTSSSESTEDERIGEDTPELGIQNPPLAASLEATPAFLADSRTPAGRF 580
624 STQERIQARLSSVIANQDPIAV 645
581 STPEELQARLSSVIANQDPIAV 602

RESULT 9
Q6DR99 ID Q6DR99 PRELIMINARY; PRT; 700 AA.
AC Q6DR99;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Neuregulin-1 type III betal-a.
OS Name=Nrg1;
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57/BL6J;
RA Anton E.S., Ghasghaei H.T., Weber J.L., McCann C., Fischer T.M.,
RA Cheung I.D., Gassmann M., Messing A., Klein R., Schwab M.H.,
RA Lloyd K.C., Lai C.;
RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
DR EMBL; AY648975; AAT68240.1;
DR GO; GO:0005102; F:receptor binding; IEA.
DR GO; GO:0009790; P:embryonic development; IEA.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002154; Neuregulin.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF02158; Neuregulin; 1.
DR PRINTS; PR01089; NEUREGULIN.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS50026; EGF_3; 1.
KW EGF-like domain.
SQ SEQUENCE 700 AA; 76504 MW; 37D7928FD7D9AC9 CRC64;

Query Match 71.6%; Score 2395; DB 2; Length 700;
Best Local Similarity 94.1%; Pred. No. 2e-123;
Matches 450; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

QY 168 STSTSTGTHLVKCAEKEKTCFVNGGECFVWKDLSNPSRYLCKPNBFTGDRCONYMA 227
DB 223 STSTSTGTHLVKCAEKEKTCFVNGGECFVWKDLSNPSRYLCKPNBFTGDRCONYMA 282
QY 228 SFYKHLGIEFMEAEELYQKRVLTITGICALLVVGIMCVAYCKTKOKKHLDRSLRSL 287
DB 283 SFYKHLGIEFMEAEELYQKRVLTITGICALLVVGIMCVAYCKTKOKKHLDRSLRSL 342

```


QY 288 RSRNNMNIANGPHHPNPPENPVOLVNOYVSKNVISSHHIVERAEATSFSTSHYTSTA 347
DB 343 RSRNNMNIANGPHHPNPPENPVOLVNOYVSKNVISSHHIVERAEATSFSTSHYTSTA 402
QY 348 HSTTVTQTPSHWSNGHTSILSESHSVTMSVNSRSHSSPTGGPRLNGTGGPRECN 407
DB 403 HSTTVTQTPSHWSNGHTSILSESHSVTMSVNSRSHSSPTGGPRLNGTGGPRECN 462
QY 408 SFRLHARETPDSYRSPHSRYSVAMTTPARMSPVDFHTPPSPKPPSEMPPVSSMTVS 467
DB 463 SFRLHARETPDSYRSPHSRYSVAMTTPARMSPVDFHTPPSPKPPSEMPPVSSMTVS 522
QY 468 MFSMAVSPWNERPPLLLVTPRLREKKPDHHPQFSSFHHPNPAHNSLSPASPLRIVED 527
DB 523 MFSMAVSPWNERPPLLLVTPRLREKKPDHHPQFSSFHHPNPAHNSLSPASPLRIVED 582
QY 528 EYETTOEYEPQEPVKLANRRRAKRTKPNGHIANRLVDSTSSQSSNSESETERV 587
DB 583 EYETTOEYEPQEPVKLANRRRAKRTKPNGHIANRLVDSTSSQSSNSESETERV 642
QY 588 GEDTFLGIONPLAASLEATPAFLADSTNPAGRFSTOEIOEARLSSVIANQDPIAV 645
DB 643 GEDTFLGIONPLAASLEATPAFLADSTNPAGRFSTOEIOEARLSSVIANQDPIAV 700

RESULT 10
Q9ESB1 PRELIMINARY; PRT; 700 AA.
AC Q9ESB1; (1)
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE SMDF neuregulin beta la.
GN Name=Nrg1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Carroll S.L., Anderson K.D., Frohnert P.W.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC EMBL; AF194438; AAG28427.1; -.
DR HSSP; Q12780; 1HRE.
DR GO; GO:0005102; F:receptor binding; IEA.
DR GO; GO:0009790; P:embryonic development; IEA.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR002114; HPr_Serp_S.
DR InterPro; IPR006210; IEGF.
DR Pfam; PF02158; Neuregulin; 1.
DR Pfam; PF02158; Neuregulin; 1.
DR PRINTS; PR01089; NEUREGULIN.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
KW EGF-like domain.
SQ SEQUENCE 700 AA; 76386 MW; 2F8111B17ECC49DA CRC64;

Query Match 70.38; Score 2354.5; DB 2; Length 700;
Best Local Similarity 80.68; Pred. No. 3.3e-121;
Matches 456; Conservative 29; Mismatches 52; Indels 29; Gaps 4;

QY 107 SGEYMKVSKLGNDSANITVESNEIITGMPA-----STEGAYVSS 150
DB 137 SSEAYTSPVSK-AQSEAGAHVTVQGDHAAVASEPSAVPTRKNRLSAPPPHSTAPPFSP 195
QY 151 ESPRISVSTEGANTSS-----STSTSTTGTSHLVKCAEKEKTCVNGGECFV 199

DB 196 ARTPEVTPKSGTQPQTETNLQTPAKLSTSTSTTGTSHLIKCAEKEKTCVNGGECFV 255
QY 200 KDLSPSRYLCKPNFTGDRCONYMASFYKHLGIEFMEAEELYOKRVLTITGICALL 259
DB 256 KDLSPSRYLCKPNFTGDRCONYMASFYKHLGIEFMEAEELYOKRVLTITGICALL 315
QY 260 VVGIMCVVAYCKTKKORLHDLRSLRSERNMNIANGPHHPNPPENPVOLVNOYVS 319
DB 316 VVGIMCVVAYCKTKKORLHDLRSLRSERNMNIANGPHHPNPPENPVOLVNOYVS 375
QY 320 KNVISSEHIVERAEATSFSTSHYTSTAHSHTTVTQTPSHWSNGHTSILSESHSVTMS 379
DB 376 KNVISSEHIVERAEATSFSTSHYTSTAHSHTTVTQTPSHWSNGHTSILSESHSVTMS 435
QY 380 SVNSRSHSSPTGGPRLNGTGGPRECNFLRHARETPDSYRSPHSRYSVAMTTPARM 439
DB 436 SVNSRSHSSPTGGPRLNGTGGPRECNFLRHARETPDSYRSPHSRYSVAMTTPARM 494
QY 440 SPVDFHTPPSPKPPSEMSPVSMVSPMSVAVSPFVEERPLLLVTPRLREKKPDH 499
DB 495 SPVDFHTPPSPKPPSEMSPVSMVSPMSVAVSPFVEERPLLLVTPRLREKKPDH 554
QY 500 PQQSSFHHPNPAHNSLSPASPLRIVEDDEYETTOEYEPQEPVKLANRRRAKRTKPN 559
DB 555 PQQSSFHHPNPAHNSLSPASPLRIVEDDEYETTOEYEPQEPVKLANRRRAKRTKPN 614
QY 560 HIANRLVDSTSSQSSNSESETERVGDTPFLGIONPLAASLEATPAFLADSTN 619
DB 615 HIANRLVDSTSSQSSNSESETERVGDTPFLGIONPLAASLEATPAFLADSTN 674
QY 620 AGFSTOEIOEARLSSVIANQDPIAV 645
DB 675 AGFSTOEIOEARLSSVIANQDPIAV 700

RESULT 11
Q9ESB0 PRELIMINARY; PRT; 695 AA.
AC Q9ESB0; (1)
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE SMDF neuregulin alpha 2a.
GN Name=Nrg1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=BDIX;
RA Carroll S.L., Anderson K.D., Frohnert P.W.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC EMBL; AF194439; AAG28428.1; -.
DR HSSP; Q12780; 1HRE.
DR GO; GO:0005102; F:receptor binding; IEA.
DR GO; GO:0009790; P:embryonic development; IEA.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR002114; HPr_Serp_S.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002154; Neuregulin.
DR Pfam; PF02158; Neuregulin; 1.
DR Pfam; PF02158; Neuregulin; 1.
DR PRINTS; PR01089; NEUREGULIN.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00026; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
KW EGF-like domain.
SQ SEQUENCE 695 AA; 75646 MW; 5277F2CBA2FB6878 CRC64;

Query Match 66.1%; Score 2211; DB 2; Length 695;
 Best Local Similarity 77.6%; Pred. No. 2.5e-113;
 Matches 440; Conservative 33; Mismatches 58; Indels 36; Gaps 8;

QY 107 SGEYMKVYKLGNDASANITIVENEITGPAS--TEGAVVSS-----ESP 153
 DB 137 SSEAYTSPVSK-AQSEAGAHVTVQGDHAAVASEPSAVPTRKNLSAFPFPFHTAPFPSP 195
 QY 154 IRI-SYSTEGANTSS-----STSTSTTGTSHLVKCAEKEKTCVNGGECFV 199
 DB 196 ARTPEVTRPKSGTOPOTETNLTQAPKLSSTSTTGTSHLVKCAEKEKTCVNGGECFV 255
 QY 200 KDLNPNRSLVCKPCNFTGDRQNYWASPYKHLGHEPME-ABELYQKRVLTITGICIAL 258
 DB 256 KDLNPNRSLVCKPCNFTGDRQNYWASPYKHLGHEPME-ABELYQKRVLTITGICIAL 309
 QY 259 LVVGIMCVAYCKTKQKQLHDLRLQSLRSERNMMNTANGPHHNPENNVQV 318
 DB 310 LVVGIMCVAYCKTKQKQLHDLRLQSLRSERNMMNTANGPHHNPENNVQV 369
 QY 319 SKNVISEHIVERAEATSTSTSHYTTAHTSTTTPSHWSNGHTEISLSESHSVIWM 378
 DB 370 SKNVISEHIVERAEATSTSTSHYTTAHTSTTTPSHWSNGHTEISLSESHSVIWM 429
 QY 379 SSVENSRHSPTGCGRLNGTGGPCRCNSFLRHARETSDSYRDSPHSRYVSAMTTPAR 438
 DB 430 SSVENSRHSPTGCGRLNGTGGPCRCNSFLRHARETSDSYRDSPHSRYVSAMTTPAR 488
 QY 439 MSPVDFTHTSSPKSPSEMSPPVSSMTVSPMSVAVSPFMEERPLLVTTPRLREKFPD 498
 DB 489 MSPVDFTHTSSPKSPSEMSPPVSSMTVSPMSVAVSPFMEERPLLVTTPRLREKFPD 548
 QY 499 HPQOFSFHNPAHNSLPSPLRIVEDEEYETQYEPAPQVKKLANSRAKETKN 558
 DB 549 HPQOFSFHNPAHNSLPSPLRIVEDEEYETQYEPAPQVKKLANSRAKETKN 608
 QY 559 GHIANRLEVDNSTSSQSSNSESETEDEYVGDGPPFLGIGNQNLAAASLEATPAFLADSRN 618
 DB 609 GHIANRLEVDNSTSSQSSNSESETEDEYVGDGPPFLGIGNQNLAAASLEATPAFLADSRN 668
 QY 619 PAGRFSTQBIQARLSSVTANQDPIAV 645
 DB 669 PAGRFSTQBIQARLSSVTANQDPIAV 695

RESULT 12

ID NRGI_XENLA STANDARD; PRT; 677 AA.
 AC Q93383; Q9W6N0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Pro-neuregulin-1 precursor (Pro-NRGI) [Contains: Neuregulin-1].
 GN Names=NRGI;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RN SEQUENCE FROM N.A. (ISOFORM ALPHA1), AND ALTERNATIVE SPLICING.
 RX MEDLINE=98352126; PubMed=9685585; DOI=10.1016/S0169-328X(98)00085-0;
 RA Yang J.F., Zhou H., Pun S., Ip N.Y., Peng H.B., Tsai K.W.K.;
 RT "Cloning of cDNAs encoding xenopus neuregulin: expression in myotomal
 muscle during embryo development."
 RL Brain Res. Mol. Brain Res. 58:59-73(1998).
 RN [2]
 RN SEQUENCE FROM N.A. (ISOFORM CRD).
 RX MEDLINE=99316087; PubMed=10383827; DOI=10.1006/mcne.1999.0759;
 RA Yang J.F., Zhou H., Choi R.C., Ip N.Y., Peng H.B., Tsai K.W.K.;
 RT "A cysteine-rich form of Xenopus neuregulin induces the expression of
 acetylcholine receptors in cultured myotubes.";

RL Mol. Cell. Neurosci. 13:415-429(1999).
 CC -!- FUNCTION: Direct ligand for the ERBB tyrosine kinase receptors.
 CC Induces expression of acetylcholine receptor in synaptic nuclei.
 CC -!- SUBCELLULAR LOCATION: Exists as a type I membrane protein and as a
 CC proteolytically released soluble growth factor form. The membrane-
 CC bound form does not seem to be active (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Comment=Additional isoforms seem to exist. Isoforms have alpha-
 CC or beta-type EGF-like domains;
 CC Name=Alpha1;
 CC IsoId=O93383-1; Sequence=Displayed;
 CC Name=CRD; Synonyms=CRD-NRGI, Cysteine-rich domain;
 CC IsoId=O93383-2; Sequence=VSP 003449, VSP 003450;
 CC -!- TISSUE SPECIFICITY: Isoform alpha1 is expressed in brain and
 CC muscle. Isoform CRD is expressed in brain and spinal cord, but at
 CC very low level in muscle.
 CC -!- DEVELOPMENTAL STAGE: Strong expression in developing brain and
 CC spinal cord of the embryo. Also expressed in the myotomal muscle.
 CC -!- DOMAIN: The cytoplasmic domain may be involved in the regulation
 CC of trafficking and proteolytic processing. Regulation of the
 CC proteolytic processing involves initial intracellular domain
 CC dimerization (By similarity).
 CC -!- DOMAIN: ERBB receptor binding is elicited entirely by the EGF-like
 CC domain.
 CC -!- PTM: Proteolytic cleavage close to the plasma membrane on the
 CC external face leads to the release of the soluble growth factor
 CC form.
 CC -!- PTM: Extensive glycosylation precedes the proteolytic cleavage (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to the neuregulin family.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF076618; AAC26804.1; -;
 CC EMBL; AF142632; AAD33893.1; -;
 CC HSP; O12780; 1HRE.
 CC InterPro; IPR000742; EGF 2.
 CC InterPro; IPR006209; EGF-like.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR002154; Neuregulin.
 CC Pfam; PF00008; EGF; 1.
 CC Pfam; PF00047; ig; 1.
 CC Pfam; PF02158; Neuregulin; 1.
 CC PRINTS; PRO1089; NEUREGULIN.
 CC PROSITE; PS00022; EGF 1; 1.
 CC PROSITE; PS01186; EGF 2; 1.
 CC PROSITE; PS00026; EGF 3; 1.
 CC PROSITE; PS00835; IG Like; 1.
 CC Alternative splicing; EGF-like domain; Glycoprotein; Growth factor;
 CC Immunoglobulin domain; Transmembrane.
 CC Chain 1 259
 CC Chain 1 677
 CC Chain 1 677
 CC Domain 1 260
 CC Transmem 261 280
 CC Domain 281 677
 CC Domain 37 132
 CC Domain 188 232
 CC Disulfid 57 116
 CC Disulfid 192 206
 CC Disulfid 200 220
 CC Disulfid 222 231
 CC Disulfid 231 25
 CC Carboxy 124 124
 CC N-linked (GlcNAc...) (Potential).
 CC Extracellular (Potential).
 CC Internal signal sequence (Potential).
 CC Cytoplasmic (Potential).
 CC Ig-like C2-type.
 CC EGF-like.
 CC By similarity.
 CC By similarity.
 CC By similarity.
 CC Lys-rich.
 CC N-linked (GlcNAc...) (Potential).


```

QY 421 RDSPHSERY 429
Db 416 RDSPHSERH 424

RESULT 14
Q35947
ID Q35947 PRELIMINARY; PRT; 461 AA.
AC Q35947;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Neuregulin.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RA Velasco J.A., Feijoo E., Avila M.A., Notario V.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
DR EMBL; U96612; BAB71812.1; -.
DR HSP; Q12780; IHR.
DR GO; GO:0005102; F:receptor binding; IEA.
DR GO; GO:0009790; P:embryonic development; IEA.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR002154; Neuregulin.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00047; ig; 1.
DR PRINTS; PR01089; Neuregulin; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS01186; EGF 2; 1.
DR PROSITE; PS50026; EGF 3; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW EGF-like domain.
SQ SEQUENCE 461 AA; 50890 MW; 935C9560F7148336 CRC64;

Query Match 57.2%; Score 1916; DB 2; Length 461;
Best Local Similarity 86.7%; Pred. No. 2.3e-97;
Matches 373; Conservative 19; Mismatches 30; Indels 8; Gaps 3;

QY 1 MSERKEGRGKGKGGKPPESAGSOSPALPQKEMKQESAGSKLVLRCTS 60
Db 1 MSERKEGRGKGKGGKPPESAGSOSPALPQKEMKQESAGSKLVLRCTS 60

QY 61 SEYSSLRKFWKNGNELNKNKPNQIKQKPKSELINKASLADSGYMKVSKLGN 120
Db 61 SEYPELRFKFWKNGSELNKRTPQIKQKPKSELINKASLADSGYMKVSKLGN 120

QY 121 DSANITIVESNEITGMPASTGAYVSSSEPIRISVSTEGANTSSSTSTSTGTHLV 180
Db 121 DSANITIVDSNEFITGMPASTRAYVSSSEPIRISVSTEGANTSSSTSTSTGTHLV 180

QY 181 KCAEKEKTFVNGGECFVKWDLNPSRYLCKCPNEFTGDRCONVYMASFYKHLGIEFME- 239
Db 181 KCAEKEKTFVNGGECFVKWDLNPSRYLCKCPNEFTGDRCONVYMASFYKHLGIEFME- 234

QY 240 AEELYQKRVLTITGICALLVVGIMCVAYCKTKQKQKLDRLQSLRSERNMMNIAN 299
Db 235 AEELYQKRVLTITGICALLVVGIMCVAYCKTKQKQKLDRLQSLRSERNMMNIAN 294

QY 300 GPHHPNPPENLVNQYVSKNVISSEHIVERAETSFSHTYSTTAHSTTTVTPPSHS 359
Db 295 GPHHPNPPETVLVNVQYVSKNIISSSEHIVERAETSFSHTYSTTAHSTTTVTPPSHS 354

```

```

QY 360 WNGHTESILSESHSVIVMSVSVNHSRHSPTGPRGRNGTGGPRECNFLRHARETPDS 419
Db 355 WNGHTESIVSESHSVIVMSVSVNHSRHSPTGPRGRNGTGGPRECNFLRHARETPDS 413

QY 420 YRDSPHSERY 429
Db 414 YRDSPHSERH 423

RESULT 15
Q6TGK9
ID Q6TGK9 PRELIMINARY; PRT; 394 AA.
AC Q6TGK9;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Neuregulin 1 alpha isoform (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Hendrickx J.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
DR EMBL; AY421758; AAR00250.1; -.
DR GO; GO:0005102; F:receptor binding; IEA.
DR GO; GO:0009790; P:embryonic development; IEA.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR002154; Neuregulin.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00047; ig; 1.
DR PRINTS; PR01089; Neuregulin; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS01186; EGF 2; 1.
DR PROSITE; PS50026; EGF 3; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW EGF-like domain.
FT NON_TER 1
FT NON_TER 394
SQ SEQUENCE 394 AA; 42980 MW; C183EB80927443F9 CRC64;

Query Match 56.0%; Score 1875.5; DB 2; Length 394;
Best Local Similarity 91.7%; Pred. No. 3.2e-95;
Matches 366; Conservative 6; Mismatches 22; Indels 5; Gaps 1;

QY 10 KKGKKKRGSGKKPPESAGSOSPALPQKEMKQESAGSKLVLRCTSSEYSSLRPK 69
Db 1 KKGKKKRGSGKKPPADGGPSPALPRLKEMKQESAGSKLVLRCTSSEYSSLRPK 60

QY 70 WFKNGNELNKNKPNQIKQKPKSELINKASLADSGYMKVSKLGNDSANIT 129
Db 61 WFKNGNELNKNKPNQIKQKPKSELINKASLADSGYMKVSKLGNDSANIT 120

QY 130 VESNEIITGMPASTGAYVSSSEPIRISVSTEGANTSSSTSTSTGTHLVKCAEKEKTF 189
Db 121 VESNEIITGMPASTRAYVSSSEPIRISVSTEGANTSSSTSTSTGTHLVKCAEKEKTF 180

QY 190 CVNGGECFVKWDLNPSRYLCKCPNEFTGDRCONVYMASFYKHLGIEFMEAEELYQKRV 249
Db 181 CVNGGECFVKWDLNPSRYLCKCPNEFTGDRCONVYMASFYKHLGIEFMEAEELYQKRV 235

```

Qy	250	TTIGICIA	LVGINCV	WVAYCK	TQKQK	UHDRL	QSLR	SERN	NNNN	NIANG	PHHP	PNP	309
Db	236	TTIGICIA	LVGINCV	WVAYCK	TQKQK	UHDRL	QSLR	SERN	NNNN	NIANG	PHHP	PNP	295
Qy	310	NVLNQY	YSKXNI	ISSEH	IVRE	AE	TSP	SHY	TSTA	HHST	TVTQ	TPSH	SWNG
Db	296	NVLNQY	YSKXNI	ISSEH	IVRE	AE	TSP	SHY	TSTA	HHST	TVTQ	TPSH	SWNG
Qy	370	SESHVIV	WSSV	ENSR	HSPT	GGPR	GLNG	TG	GG	PRE	CNS	408	
Db	356	SESHVIV	WSSV	ENSR	HSPT	GGPR	GLNG	TG	GG	PRE	CNS	394	

Search completed: April 12, 2005, 14:33:51
Job time : 187 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 14:26:42 ; Search time 48 Seconds
(without alignments)
1292.913 Million cell updates/sec

Title: US-10-082-747A-93
Perfect score: 3347
Sequence: 1 MSERKEGKGKGGKKGKRGSS.....QBEIQARLSSVIANQDPIAV 645
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3343	99.9	645	2 B43273	herregulin, splice
2	3280	98.0	637	2 C43273	herregulin, precursor
3	3202.5	95.7	640	2 A43273	herregulin, precursor
4	3066.5	91.6	636	2 I61718	neu differentiation
5	3058.5	91.4	662	2 I61722	neu differentiation
6	2980	89.0	639	2 I61719	neu differentiation
7	2526.5	75.5	602	2 A45769	acetylcholine rece
8	2086.5	62.3	462	2 I38404	neu differentiation
9	1751.5	52.3	350	2 I38403	neu differentiation
10	1183	35.3	241	2 D43273	herregulin, precursor
11	1098	32.8	241	2 S32359	glial growth facto
12	1059	31.6	230	2 A56210	neu differentiation
13	1028	30.7	868	2 JCS701	ErbB kinase activa
14	993	29.7	850	2 JCS700	ErbB kinase activa
15	982	29.6	860	2 JCS702	ErbB kinase activa
16	827	24.7	422	2 S32357	glial growth facto
17	788	23.5	175	2 I38408	neu differentiation
18	545.5	16.3	125	2 S62676	herregulin isoform
19	485	14.5	125	2 I38405	neu differentiation
20	351	10.5	296	2 A56943	sensory/motor neur
21	334.5	10.0	713	2 T44447	neuregulin-3 (limp
22	182.5	5.5	57	2 PC4415	ErbB kinase activa
23	181	5.4	2254	2 T62215	protein T622.14 [
24	179	5.3	3507	2 T34513	hypothetical prote
25	172	5.1	734	2 B42680	nucleolus-cytoplas
26	171	5.1	1091	1 IJCHNL	neural cell adhesi
27	170	5.1	1891	2 T13594	hypothetical prote
28	169	5.0	1920	2 T13893	gene hindsight pro
29	161	4.8	1510	2 T33100	hypothetical prote

30	159	4.8	1192	2 T18611	probable serine/th
31	157	4.7	6642	2 T29757	protein UNC-89 - C
32	156.5	4.7	279	2 T16201	hypothetical prote
33	155.5	4.6	1306	2 S25370	MSB2 protein - yea
34	151.5	4.5	955	2 E84845	probable villin 2
35	151.5	4.5	7962	2 T13846	elastic titin - hu
36	151	4.5	5327	2 T13564	microtubule-associ
37	150	4.5	1092	1 JN0635	neural cell adhesi
38	149.5	4.5	1367	2 S74285	BUD3 protein - yea
39	149	4.5	409	2 S70704	carbon catabolite
40	149	4.5	1271	2 T24008	hypothetical prote
41	147.5	4.4	699	2 I38073	nucleolar phosphop
42	147	4.4	1203	2 T17415	mycellial surface a
43	146.5	4.4	402	2 E86185	hypothetical prote
44	146.5	4.4	844	2 S61112	regulatory protein
45	146	4.4	534	2 T39903	serine-rich protei

ALIGNMENTS

RESULT 1

B43273
herregulin, splice form beta 1 - human
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: B43273; I38406
R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yansu
Science 256, 1205-1210, 1992
A>Title: Identification of herregulin, a specific activator of p185(erbB2).
A:Reference number: A43273; MUID:92271253; PMID:1350381
A:Accession: B43273
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr.
A:Molecule type: mRNA
A:Residues: 1-645 <HOL>
A:Cross-references: UNIPROT:Q02297
Riwen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.,
Mol. Cell. Biol. 14, 1909-1919, 1994
A>Title: Structural and functional aspects of the multiplicity of Neu differentiation f
A:Reference number: A56210; MUID:94158863; PMID:7509448
A:Accession: I38406
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 'A', 95-418, 'F', 420-645 <RES>
A:Cross-references: EMBL:U02328; NID:G408406; PIDN:AAA19953.1; PID:G408407
C:Genetics:
A:Gene: GDB:HGL
A:Cross-references: GDB:I32656; OMIM:142445
A:Map position: 8p22-9p11
C:Superfamily: human herregulin; EGF homology; immunoglobulin homology
C:Keywords: alternative splicing
F:182-221/Domain: EGF homology <EGF>

Query Match 99.9%; Score 3343; DB 2; Length 645;
Best Local Similarity 99.8%; Pred.No. 1.3e-183;
Matches 644; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSERKEGKGKGGKKGKRGSSQSPALPPQPKEMKQESAGSKVLRCETS	60
Db	1	MSERKEGKGKGGKKGKRGSGKPPRLPKEMKQESAGSKVLRCETS	60
Qy	61	SEYSSLRPFKWFNGNGLNKNKPQNKIQKPKGKSELINKASLADSGEYMKVKISLGN	120
Db	61	SEYSSLRPFKWFNGNGLNKNKPQNKIQKPKGKSELINKASLADSGEYMKVKISLGN	120
Qy	121	DSASANTITVESNEITGMPASTEGAYVSSSPIRISVSTEGANTSSSTSTTGTSHLV	180
Db	121	DSASANTITVESNEITGMPASTEGAYVSSSPIRISVSTEGANTSSSTSTTGTSHLV	180
Qy	181	KCAEKEKTCVNGGCFWVKDLSNPSRYLCKCPNFTGDRCONYMASFYKHLGIEFMEA	240
Db	181	KCAEKEKTCVNGGCFWVKDLSNPSRYLCKCPNFTGDRCONYMASFYKHLGIEFMEA	240

QY 241 EELYQKRVLTITGICIALLVGIMCVAYCKTKKQKLDRLQSLRSENNMMNIANG 300
DB 241 EELYQKRVLTITGICIALLVGIMCVAYCKTKKQKLDRLQSLRSENNMMNIANG 300
QY 301 PHHPNPPENVQLVNOQYVKNVISSEHIVERAETSFSHTSTTAHSTTTVQTPSHW 360
DB 301 PHHPNPPENVQLVNOQYVKNVISSEHIVERAETSFSHTSTTAHSTTTVQTPSHW 360
QY 361 SNHTESILSESHSVIVMSSVENSRRHSPTGGPRGLNGTGGPRECSFLRHARETPDSY 420
DB 361 SNHTESILSESHSVIVMSSVENSRRHSPTGGPRGLNGTGGPRECSFLRHARETPDSY 420
QY 421 RDSPHSERVVSAMTTPARMSPVDFHTPPSPKSPPEMSPVSSMTVSMPSMAVSPFMEEE 480
DB 421 RDSPHSERVVSAMTTPARMSPVDFHTPPSPKSPPEMSPVSSMTVSMPSMAVSPFMEEE 480
QY 481 RPLLLVTPPRLREKKFDHHPQFSSFFHNPADHNSLSPASPLRIVEDEEYETTQYEPAQ 540
DB 481 RPLLLVTPPRLREKKFDHHPQFSSFFHNPADHNSLSPASPLRIVEDEEYETTQYEPAQ 540
QY 541 EPVKKLANSRRAKTKPNGHIANRLVDSNTSSQSSNSESETEDERVGEDTPFLGIQNPL 600
DB 541 EPVKKLANSRRAKTKPNGHIANRLVDSNTSSQSSNSESETEDERVGEDTPFLGIQNPL 600
QY 601 AASLEATPAFLADSRTPNAGRFSTQBEIOARLSSVIANQDPIAV 645
DB 601 AASLEATPAFLADSRTPNAGRFSTQBEIOARLSSVIANQDPIAV 645
RESULT 2
C43273
heregulin precursor, splice form beta-2 - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 08-Sep-2002
C:Accession: C43273; I38407
R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yansu
Science 256, 1205-1210, 1992
A:Title: Identification of heregulin, a specific activator of p185(erbB2).
A:Reference number: A43273; MUID:92271253; PMID:1350381
A:Accession: C43273
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-637 <HOL>
R:Wen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.;
Mol. Cell. Biol. 14, 1909-1919, 1994
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation fa
A:Reference number: A56210; MUID:94158863; PMID:7509448
A:Accession: I38407
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 119-406 <RES>
A:Cross-references: EMBL:U02329; NID:G408408; PIDN:AAA19954.1; PID:G408409
C:Genetics:
A:Gene: GDB:HGL
A:Cross-references: GDB:132656; OMIM:142445
A:Map position: 8p22-8p11
C:Superfamily: human heregulin; EGF homology; immunoglobulin homology
C:Keywords: alternative splicing
F:182-221/Domain: EGF homology <EGF>
Query Match 98.0%; Score 3280; DB 2; Length 637;
Best Local Similarity 98.4%; Pred. No. 5.1e-180;
Matches 635; Conservative 1; Mismatches 1; Indels 8; Gaps 1;
QY 1 MSERKEGRGKGKKERSGKPPSAAGSQSPALPPQLKEMKSOESAAGSKVLRCETS 60
DB 1 MSERKEGRGKGKKERSGKPPSAAGSQSPALPPQLKEMKSOESAAGSKVLRCETS 60
QY 61 SYSSLSLRFKFWKNGNELNKNPNQIKIQKPKGKSELINKASLADSGEYCMCKVLSKLN 120
DB 61 SYSSLSLRFKFWKNGNELNKNPNQIKIQKPKGKSELINKASLADSGEYCMCKVLSKLN 120
QY 121 DSASANITIVESNEITGMPASTEGAYVSSSESPIRISVSTEGANTSSSTSTTTGTSHLV 180

DB 121 DSASANITIVESNEITGMPASTEGAYVSSSESPIRISVSTEGANTSSSTSTTTGTSHLV 180
QY 181 KCASEKTKFCVNGGECFMVKDLSNPRLYKCKPNEFTGDRCONTVMASFYKHLGIEPMEA 240
DB 181 KCASEKTKFCVNGGECFMVKDLSNPRLYKCKPNEFTGDRCONTVMASFYK 232
QY 241 EELYQKRVLTITGICIALLVGIMCVAYCKTKKQKLDRLQSLRSENNMMNIANG 300
DB 241 EELYQKRVLTITGICIALLVGIMCVAYCKTKKQKLDRLQSLRSENNMMNIANG 292
QY 301 PHHPNPPENVQLVNOQYVKNVISSEHIVERAETSFSHTSTTAHSTTTVQTPSHW 360
DB 301 PHHPNPPENVQLVNOQYVKNVISSEHIVERAETSFSHTSTTAHSTTTVQTPSHW 352
QY 361 SNHTESILSESHSVIVMSSVENSRRHSPTGGPRGLNGTGGPRECSFLRHARETPDSY 420
DB 361 SNHTESILSESHSVIVMSSVENSRRHSPTGGPRGLNGTGGPRECSFLRHARETPDSY 412
QY 421 RDSPHSERVVSAMTTPARMSPVDFHTPPSPKSPPEMSPVSSMTVSMPSMAVSPFMEEE 480
DB 421 RDSPHSERVVSAMTTPARMSPVDFHTPPSPKSPPEMSPVSSMTVSMPSMAVSPFMEEE 472
QY 481 RPLLLVTPPRLREKKFDHHPQFSSFFHNPADHNSLSPASPLRIVEDEEYETTQYEPAQ 540
DB 481 RPLLLVTPPRLREKKFDHHPQFSSFFHNPADHNSLSPASPLRIVEDEEYETTQYEPAQ 532
QY 541 EPVKKLANSRRAKTKPNGHIANRLVDSNTSSQSSNSESETEDERVGEDTPFLGIQNPL 600
DB 541 EPVKKLANSRRAKTKPNGHIANRLVDSNTSSQSSNSESETEDERVGEDTPFLGIQNPL 592
QY 601 AASLEATPAFLADSRTPNAGRFSTQBEIOARLSSVIANQDPIAV 645
DB 593 AASLEATPAFLADSRTPNAGRFSTQBEIOARLSSVIANQDPIAV 637
RESULT 3
C43273
heregulin precursor, splice form alpha - human
N:Alternate names: breast cancer cell differentiation factor p45; Neu differentiation fa
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 08-Sep-2002
C:Accession: A43273; A48498; A38155
R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yansu
Science 256, 1205-1210, 1992
A:Title: Identification of heregulin, a specific activator of p185(erbB2).
A:Reference number: A43273; MUID:92271253; PMID:1350381
A:Accession: A43273
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-640 <HOL>
A:Experimental source: breast tumor cell line, MDA-MB-231, ATCC HTB 26
A:Note: sequence extracted from NCBI backbone (NCBIP:103250)
R:Culouscou, J.M.; Plozman, G.D.; Carlton, G.W.; Green, J.M.; Shoyab, M.
J. Biol. Chem. 268, 18407-18410, 1993
A:Title: Characterization of a breast cancer cell differentiation factor that specifical
A:Reference number: A48498; MUID:93366731; PMID:7689552
A:Accession: A48498
A:Molecule type: protein
A:Residues: 20-21,'X',23-24,'XX',27-28 <CUL>
R:Pelles, E.; Bacus, S.S.; Koski, R.A.; Lu, H.S.; Wen, D.; Ogden, S.G.; Levy, R.B.; Yarde
Cell 69, 205-216, 1992
A:Title: Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein that induce
A:Reference number: A38155; MUID:92208945; PMID:1348215
A:Accession: A38155
A:Molecule type: protein
A:Residues: 'X',15-16,'X',18-20,'RG',23-24,'GP',27,'E',29,'XP',32-36 <PEL>
A:Note: sequence extracted from NCBI backbone (NCBIP:91347)
C:Genetics:
A:Gene: GDB:HGL
A:Cross-references: GDB:132656; OMIM:142445
A:Map position: 8p22-8p11
C:Superfamily: human heregulin; EGF homology; immunoglobulin homology

A:Accession: I61720
A:Cross-references: EMBL:U03321; NID:G408392; PIDN:AAA19946.1; PID:G408393

A;Accession: 161720
 A;Cross-references: EMBL:U02321; NID:G408392; PIDN:AAA19946.1; PID:G408393
 A;References: 1-411; R:940050, 94025

A,Cross references: Q9AR0L;P3J228; B6B0;O0024; A003400;
C,Superfamily: human heregulin; EGF homology; immunoglobulin homology
F;182-221/Domain: EGF homology <EGF>

Db 594 LAASLEVAFAFRLAESTNPAGRSTQBELQARLSSVIANQDPIAV 639

RESULT 7

A45769
C:Species: Gallus gallus (chicken)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C/Accession: A45769
R:Fallis, D.L.; Rosen, K.M.; Corfas, G.; Lane, W.S.; Fischbach, G.D.
Cell 72, 801-815, 1993
A>Title: ARIA, a protein that stimulates acetylcholine receptor synthesis, is a member of
A/Reference number: A45769; MUID:93201602; PMID:8453670
A/Accession: A45769
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-602 <PAL>
A/Cross-references: UNIPROT:Q05199; GB:L11264; NID:g212603; PIDN:AAA49037.1; PID:g212604
A/Experimental source: brain
C:Superfamily: human heregulin; EGF homology; immunoglobulin homology

Query Match 75.5%; Score 2526.5; DB 2; Length 602;
Best Local Similarity 77.2%; Pred. No. 5.3e-137; Indels 37; Gaps 4;
Matches 480; Conservative 54; Mismatches 51;
QY 25 ESAAGSQSPALPQKEMKQSAAGSKLVLCRTSESYSLRFKQNGNELNRKPKQ 84
DB 17 QTDVNSSYNTVPKLEKMQEAVGVQKLVLCRTSEYPALEKQNGKEITKQNRPE 76
QY 85 NIKIQKPGK-SLEIRNKASLADSGEYMKVSKLGNDSANITIVESNEIITGMPAST 143
DB 77 NVKIPKQKYSLEHYRATLADAGEYACRVSKLGNDSKASVIITDTN----- 126
QY 144 EGAVSSESPIRISVSTEGANTSSSTSTSTGTSNHLVCAEKEKTCVNGGECFVMDLS 203
DB 127 -----ATSTSTGTSNHLVCAEKEKTCVNGGECFVMDLP 162
QY 204 NPSRYLCKPNEFTGDRCONVYMASFYKHLGIEFMEAEELYQKRVLTITGICALLVVG 263
DB 163 NPPRYLCRCNEFTGDRCONVYMASFYKHLGIEFMEAEELYQKRVLTITGICALLVVG 222
QY 264 MCVAICYKQKQKLDRLRQSLRERNMMNIANGPHHPNPPNPNVQVYVSKNVI 323
DB 223 MCVAICYKQKQKLDRLRQSLRERNMMNIANGPHHPNPPNPNVQVYVSKNII 282
QY 324 SSEHIVERAETSFSHTSTAHSTTTVOTPSHWSNGHTESILSESHVIMSVSN 383
DB 283 SSERVVERETSFSTSHYTSITTHSMVTVTQTPSHWSNGHTESILSESHVIMSVSN 342
QY 384 SRHSSPTGGPRGLNGTGGPRECNSFLRHARETPDSYRSPHSERYVSAMTTPARMSPVD 443
DB 343 SRHTSPT-GPRGLNGIGGPRECNSFLRHARETPDSYRSPHSERYVSAMTTPARMSPVD 401
QY 444 FHTPSPGKPPSPMSPPVSMVMSMAVSPVMEERPLLLVTPRLREKFDHPHQPF 503
DB 402 FHTPSPGKPPSPMSPPVSSLTISPSVAVSPVMEERPLLLVTPRLRE-KYDNLHQF 460
QY 504 SSEHNPANDSNLSPASPLRIVEDEEYETOEYEPAPQPKLANSPRAKTPKNGHIAN 563
DB 461 NSPHNPHTESNLSPSPURIVEDEEYETOEYEPAPQPKLTNSRRVKRTPKNGHISS 520
QY 564 RLEVDNNTSQSSNSESETEDEVRGDTPLGQNLAAASLEATPAFLADSRNTPAGRF 623
DB 521 RVEVDSDTSQSSTSESETEDEIRGDTPLSLQNPWATSLPAAAYRLAENRTNPANRF 580
QY 624 STQEEIQAARLSSVIANQDPIAV 645
DB 581 STPEELQAARLSSVIANQDPIAV 602

RESULT 8

138404

neu differentiation factor - human

C:Species: Homo sapiens (man)
C/Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C/Accession: I38404
R:Wen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.; Mol. Cell. Biol. 14, 1909-1919, 1994
A/Title: Structural and functional aspects of the multiplicity of Neu differentiation f
A/Reference number: A56210; MUID:94158863; PMID:7509448
A/Accession: I38404
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-462 <RES>
A/Cross-references: UNIPROT:Q02297; EMBL:U02326; NID:g408402; PIDN:AAA19951.1; PID:g408
C:Superfamily: human heregulin; EGF homology; immunoglobulin homology

Query Match 62.3%; Score 2086.5; DB 2; Length 462;
Best Local Similarity 94.6%; Pred. No. 5.2e-112; Indels 5; Gaps 1;
Matches 406; Conservative 3; Mismatches 15;
QY 1 MSERKEGRGKGGKKGKSGKPPSAAGSQSPALPQKEMKQSAAGSKLVLCRTS 60
DB 1 MSERKEGRGKGGKKGKSGKPPSAAGSQSPALPQKEMKQSAAGSKLVLCRTS 60
QY 61 SEYSSLRFPKFWKNGNELNRKPKQNIQKQKPKSELINKASLADSGEYMKVSKLGN 120
DB 61 SEYSSLRFPKFWKNGNELNRKPKQNIQKQKPKSELINKASLADSGEYMKVSKLGN 120
QY 121 DSANITIVESNEIITGMPASTGAYVSSSEPIRISVSTEGANTSSSTSTGTSNHLV 180
DB 121 DSANITIVESNEIITGMPASTGAYVSSSEPIRISVSTEGANTSSSTSTGTSNHLV 180
QY 181 KCAEKEKTCVNGGECFVMDLSNPSRYLCKPNEFTGDRCONVYMASFYKHLGIEFME 240
DB 181 KCAEKEKTCVNGGECFVMDLSNPSRYLCKPNEFTGDRCONVYMASFYKHLGIEFME 235
QY 241 ELYQKRVLTITGICALLVVGIMCVAYCKTKQKQKLDRLRQSLRERNMMNIANG 300
DB 236 ELYQKRVLTITGICALLVVGIMCVAYCKTKQKQKLDRLRQSLRERNMMNIANG 295
QY 301 PHHPNPPNPNVQVYVSKNVISSSEHIVERAETSFSHTSTAHSTTTVOTPSHWS 360
DB 296 PHHPNPPNPNVQVYVSKNVISSSEHIVERAETSFSHTSTAHSTTTVOTPSHWS 355
QY 361 SNGHTESILSESHVIMSVSNRSHSSPTGGPRGLNGTGGPRECNSFLRHARETPDSY 420
DB 356 SNGHTESILSESHVIMSVSNRSHSSPTGGPRGLNGTGGPRECNSFLRHARETPDSY 415
QY 421 RDSPHSERY 429
DB 416 RDSPHSERH 424

RESULT 9

I38403
neu differentiation factor - human (fragment)
C:Species: Homo sapiens (man)
C/Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 08-Sep-2002
C/Accession: I38403
R:Wen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.; Mol. Cell. Biol. 14, 1909-1919, 1994
A/Title: Structural and functional aspects of the multiplicity of Neu differentiation f
A/Reference number: A56210; MUID:94158863; PMID:7509448
A/Accession: I38403
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-350 <RES>
A/Cross-references: EMBL:U02325; NID:g408400; PIDN:AAA19950.1; PID:g408401
C:Superfamily: human heregulin; EGF homology; immunoglobulin homology

Query Match 52.3%; Score 1751.5; DB 2; Length 350;
Best Local Similarity 95.1%; Pred. No. 5.1e-93; Indels 3; Gaps 1;
Matches 333; Conservative 4; Mismatches 10;

QY 194 GECFMVKDLSNPSRYLCKPNEFTGDRCONVY----MASFYKHLGIFMEABELYQKRVLT 250
Db 1 GECFMVKDLSNPSRYLCKPNEFTGDRCONVY----MASFYKHLGIFMEABELYQKRVLT 60
QY 251 ITGICITALLVVGIMCVAVYCKTKQKRLHDLRQSLRSENNMMNIANGPHHPNPPEN 310
Db 61 ITGICITALLVVGIMCVAVYCKTKQKRLHDLRQSLRSENNMMNIANGPHHPNPPEN 120
QY 311 VOLVNOYVSKNVISSHHIVERAETSFSHTYSTAHSTTQTTPSHSWSNGHTESILS 370
Db 121 VOLVNOYVSKNVISSHHIVERAETSFSHTYSTAHSTTQTTPSHSWSNGHTESILS 180
QY 371 ESHSVIWMSSVENSRSSTGPRGLNGTGPRECNSFLRHARETDSYRDSPHSERV 430
Db 181 ESHSVIWMSSVENSRSSTGPRGLNGTGPRECNSFLRHARETDSYRDSPHSERV 240
QY 431 SAMTTPARMSVDFTHTPSPKPPSPMSPPVSMVTVMSPMAVSPPMESERPLLVTTPR 490
Db 241 SAMTTPARMSVDFTHTPSPKPPSPMSPPVSMVTVMSPMAVSPPMESERPLLVTTPR 300
QY 491 LREKKFDHPQOFSFPHNPAHDSNLSPLRIVEDEYETTOYEPAP 540
Db 301 LREKKFDHPQOFSFPHNPAHDSNLSPLRIVEDEYETTOYEPAP 350
RESULT 10
D43273
heregulin precursor, splice form beta-3 - human
N;Alternate names: glial growth factor HRG-beta-3; heregulin
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: D43273; S32358
R;Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yamauchi
Science 256, 1205-1210, 1992
A;Title: Identification of heregulin, a specific activator of p185(erbB2).
A;Reference number: A43273; MUID:92271253; PMID:1350381
A;Accession: D43273
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A;Molecule type: mRNA
A;Residues: 1-241 <HOL>
A;Cross-references: UNIPROT:Q02297
R;Marchionni, M.A.; Goodearl, A.D.J.; Chen, M.S.; Bermingham-McDonogh, O.; Kirk, C.; Hen
les, I.; Davis, J.B.; Hsuan, J.J.; Totty, N.F.; Otsu, M.; McBurney, R.N.; Waterfield, M.
Nature 362, 312-318, 1993
A;Title: Glial growth factors are alternatively spliced erbB2 ligands expressed in the n
A;Reference number: S32357; MUID:93205115; PMID:8096067
A;Accession: S32358
A;Molecule type: mRNA
A;Residues: 1-241 <MAR>
A;Cross-references: UNIPROT:Q07112; GB:L12261; NID:9292049; PIDN:AAB59358.1; PID:g292050
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 08-Sep-2002
C;Accession: A56210
R;Wen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janesen, A.M.;
Mol. Cell. Biol. 14, 1909-1919, 1994
A;Title: Structural and functional aspects of the multiplicity of Neu differentiation fa
C;Keywords: alternative splicing
F;182-221/Domain: EGF homology <EGF>

Query Match 35.3%; Score 1183; DB 2; Length 241;
Best Local Similarity 99.6%; Pred. No. 9.6e-61;
Matches 229; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSERKEGRGKGKKGKRGSGKPPSAAGSQSPALPPOLKEMKQESAGSKLVLRCTSS 60
Db 1 MSERKEGRGKGKKGKRGSGKPPSAAGSQSPALPPOLKEMKQESAGSKLVLRCTSS 60
QY 61 SEYSSLRPFKFWKNGNELNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGN 120
Db 61 SEYSSLRPFKFWKNGNELNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGN 120
QY 121 DSANANITIVESNEITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTTGTSHLV 180
Db 121 DSANANITIVESNEITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTTGTSHLV 180
Query Match 32.8%; Score 1098; DB 2; Length 241;
Best Local Similarity 91.7%; Pred. No. 6.9e-56;
Matches 211; Conservative 9; Mismatches 10; Indels 0; Gaps 0;
QY 1 MSERKEGRGKGKKGKRGSGKPPSAAGSQSPALPPOLKEMKQESAGSKLVLRCTSS 60
Db 1 MSERKEGRGKGKKGKRGSGKPPSAAGSQSPALPPOLKEMKQESAGSKLVLRCTSS 60
QY 61 SEYSSLRPFKFWKNGNELNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGN 120
Db 61 SEYSSLRPFKFWKNGNELNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGN 120
QY 121 DSANANITIVESNEITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTTGTSHLV 180
Db 121 DSANANITIVESNEITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTTGTSHLV 180
Query Match 31.6%; Score 1059; DB 2; Length 230;
Best Local Similarity 93.6%; Pred. No. 1.1e-53;
Matches 205; Conservative 4; Mismatches 10; Indels 0; Gaps 0;
QY 12 KGKKERGGKGGKKGKRGSGKPPSAAGSQSPALPPOLKEMKQESAGSKLVLRCTSSSEYSSLRPFKWF 71
Db 1 KGKKERGGKGGKKGKRGSGKPPSAAGSQSPALPPOLKEMKQESAGSKLVLRCTSSSEYSSLRPFKWF 60
QY 72 KNGNELNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGNDSASANTTIVE 131
Db 72 KNGNELNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGNDSASANTTIVE 131

QY 181 KCAEKEKTCVNGGECFMVKDLSNPSRYLCKPNEFTGDRCONVYMASFY 230
Db 181 KCAEKEKTCVNGGECFMVKDLSNPSRYLCKPNEFTGDRCONVYMASFY 230
RESULT 11
S32359
glial growth factor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 19-Mar-1997 #sequence_revision 01-Aug-1997 #text_change 09-Jul-2004
C;Accession: S32359
R;Marchionni, M.A.; Goodearl, A.D.J.; Chen, M.S.; Bermingham-McDonogh, O.; Kirk, C.; Hen
les, I.; Davis, J.B.; Hsuan, J.J.; Totty, N.F.; Otsu, M.; McBurney, R.N.; Waterfield, M.
Nature 362, 312-318, 1993
A;Title: Glial growth factors are alternatively spliced erbB2 ligands expressed in the n
A;Reference number: S32357; MUID:93205115; PMID:8096067
A;Accession: S32359
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-241 <MAR>
A;Cross-references: UNIPROT:Q07112; GB:L12259; NID:g289413; PIDN:AAA30540.1; PID:g289414
C;Superfamily: human heregulin; EGF homology; immunoglobulin homology
F;182-221/Domain: EGF homology <EGF>

Query Match 32.8%; Score 1098; DB 2; Length 241;
Best Local Similarity 91.7%; Pred. No. 6.9e-56;
Matches 211; Conservative 9; Mismatches 10; Indels 0; Gaps 0;
QY 1 MSERKEGRGKGKKGKRGSGKPPSAAGSQSPALPPOLKEMKQESAGSKLVLRCTSS 60
Db 1 MSERKEGRGKGKKGKRGSGKPPSAAGSQSPALPPOLKEMKQESAGSKLVLRCTSS 60
QY 61 SEYSSLRPFKFWKNGNELNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGN 120
Db 61 SEYSSLRPFKFWKNGNELNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGN 120
QY 121 DSANANITIVESNEITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTTGTSHLV 180
Db 121 DSANANITIVESNEITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTTGTSHLV 180
QY 181 KCAEKEKTCVNGGECFMVKDLSNPSRYLCKPNEFTGDRCONVYMASFY 230
Db 181 KCAEKEKTCVNGGECFMVKDLSNPSRYLCKPNEFTGDRCONVYMASFY 230

RESULT 12
A56210
neu differentiation factor - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 08-Sep-2002
C;Accession: A56210
R;Wen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janesen, A.M.;
Mol. Cell. Biol. 14, 1909-1919, 1994
A;Title: Structural and functional aspects of the multiplicity of Neu differentiation fa
C;Reference number: A56210; MUID:94158663; PMID:7509448
A;Accession: A56210
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-230 <RES>
A;Cross-references: EMBL:U02315; NID:g408380; PIDN:AAA19940.1; PID:g408381
C;Superfamily: human heregulin; EGF homology; immunoglobulin homology

Query Match 31.6%; Score 1059; DB 2; Length 230;
Best Local Similarity 93.6%; Pred. No. 1.1e-53;
Matches 205; Conservative 4; Mismatches 10; Indels 0; Gaps 0;
QY 12 KGKKERGGKGGKKGKRGSGKPPSAAGSQSPALPPOLKEMKQESAGSKLVLRCTSSSEYSSLRPFKWF 71
Db 1 KGKKERGGKGGKKGKRGSGKPPSAAGSQSPALPPOLKEMKQESAGSKLVLRCTSSSEYSSLRPFKWF 60
QY 72 KNGNELNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGNDSASANTTIVE 131
Db 72 KNGNELNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGNDSASANTTIVE 131

